

UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

DEPARTAMENTO DE ESTADÍSTICA E INFORMÁTICA
FACULTAD DE ECONOMÍA Y PLANIFICACIÓN



TRABAJO INTEGRADOR

Inferencia Estadística 2025-II

Integrante	Código
Castillo Ruiz Mauricio Gabriel	20230384
Gómez Vigo Héctor Estefano	20230397
Montúfar Paiva Yeraldí Mercedes	20230400
Paucar Arango Marcos David Alexander	20221412
Rojas Taco Fabiana	20220956
Villanueva Huamani Alexander Ruben	20230419
Zavala Malpartida Kay Daniela L.	20230420

Docente: Fernando Miranda Villagómez

LA MOLINA - LIMA - PERÚ 2025

Contents

Distribución Pareto:	1
Pregunta 1	1
1.1 EMV para parametro alfa:	1
1.2 EMV para parametro Beta:	4
1.3 Estimador para la Media Poblacional (μ)	10
1.4 Método de Momentos para los parámetros de α y β de la Distribución Pareto	12
1.5 Estimador : Estimador Insesgado de Varianza Uniformemente Mínima (EIVUM) para el parámetro β	30
2) Pregunta 2 - Intervalos de confianza	35
2.1 Estimador 1 (el mínimo para alfa)	35
2.2 Estimador 2: Suma Log de T	36
2.3 Estimador 3: Media Muestral (\bar{X}) para la media poblacional μ	38
2.4 Estimador 4 : para α y β	39
2.5 Estimador 5 (para β con el EIVUM)	48
3) Pregunta 3: Prueba de Hipótesis con Datos Pareados	50
3.1 Simulación de una segunda variable correlacionada	50
3.2 Prueba de McNemar	51
3.3 Intervalo de confianza para la diferencia de proporciones	54
3.4 Análisis de Potencia de tamaño muestral	55
4) Pregunta 4 - Pruebas de hipótesis para los estimadores	57
4.1 Prueba de hipótesis para el parámetro α (EMV)	57
4.2 Prueba de Hipótesis para el parámetro β (EMV)	59
4.3 Prueba de Hipótesis para la Media Poblacional (μ)	60
4.4 Prueba de Hipótesis mediante el estimador de momentos α y β	62
4.5 Prueba de Hipótesis para β con el EIVUM	68
6. Pregunta 6 - Aplicación de una prueba de Distribución asintótica de la razón de verosimilitud.	71
6.1 Justificación del parámetro elegido	71
6.2 Formulación de hipótesis	72
6.3 Función de verosimilitud	72
6.4 Estimadores de Máxima Verosimilitud	72
6.5 Razón de Verosimilitud	72
6.6 Distribución asintótica (Teorema de Wilks)	73

6.7 P-valor y Regla de decisión	73
6.8 Conclusiones de la prueba	74

Distribución Pareto:

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}} I_{(\alpha, \infty)}(x)$$

$$E(X) = \frac{\alpha\beta}{\beta-1} ; \quad \text{VAR}(X) = \left(\frac{\alpha}{\beta-1}\right)^2 \frac{\beta}{\beta-2}$$

Pregunta 1

1.1 EMV para parametro alfa:

Usaremos el Método de máxima verosimilitud para poder hallar el primer estimador para alfa:

$$L(\alpha, \beta) = \prod_{i=1}^n f(x_i; \alpha, \beta) = \prod_{i=1}^n \frac{\beta\alpha^\beta}{x_i^{\beta+1}} I_{[\alpha, \infty)}(x_i) = \beta^n \alpha^{n\beta} \left(\prod_{i=1}^n x_i^{-(\beta+1)} \right) \times \prod_{i=1}^n I_{[\alpha, \infty)}(x_i)$$

Analizando la indicadora:

$$\begin{aligned} \prod_{i=1}^n I_{[\alpha, \infty]}(x_i) &= I\left(\bigcap_{i=1}^n \{x_i \geq \alpha\}\right) \\ &= I(\alpha \leq x_1, \alpha \leq x_2, \dots, \alpha \leq x_n) \\ &= I(\alpha \leq \min\{x_1, \dots, x_n\}) = I(\alpha \leq y_1) \end{aligned}$$

donde $y_1 = X(1)$ es el mínimo de la muestra:

El EMV de $\theta = \alpha$ es $T = Y_1$.

1.1.2 Propiedades:

a) **Insesgamiento:** Distribución del mínimo:

$$\begin{aligned} X_1 \dots X_n &\sim \text{Pareto } (\alpha, \beta) \text{ i i d} \\ F(x) &= 1 - \left(\frac{\alpha}{x}\right)^\beta, x \geq \alpha \\ y &= x_i \\ Fx_i(y) &= n[1 - F(y)]^{n-1} \mathcal{F}(y) \rightarrow n \left[\left(\frac{\alpha}{y}\right)^\beta \right]^{\beta-1} \cdot \frac{\beta\alpha^\beta}{y^{\beta+1}} \\ &= n \left[\frac{\alpha^{\beta(n-1)}}{\alpha^{\beta(n-1)}} \right] \cdot \frac{\beta\alpha^\beta}{y^{\beta+1}} = \frac{(n\beta)\alpha^{n\beta}}{y^{n\beta+1}}, y \geq \alpha \\ &\rightarrow x_1 \sim \text{Pareto}(\alpha, n\beta) \end{aligned}$$

Hay sesgo positivo, es decir que sobreestima al parámetro, no cumple propiedades de Insesgamiento. Ahora lo veremos mejor con los valores para nuestra data:

```

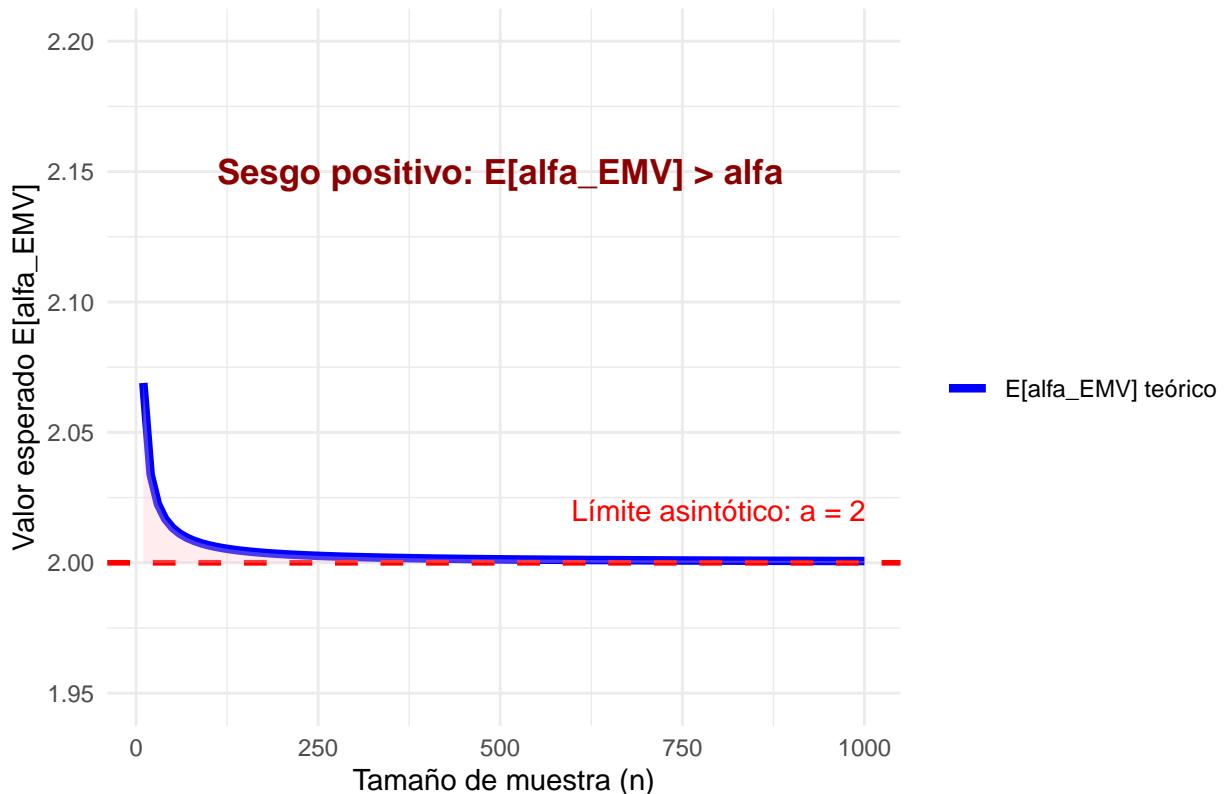
# 1. Valor observado del mínimo
alpha_EMV_observado <- min(reclamos)
# 2. Valor esperado teórico del mínimo
E_alpha_EMV_teorico <- (n * beta_true * alpha_true) / (n * beta_true - 1)

```

```
## [1] "El valor observado es: 2.00031032898679 y el valor esperado es: 2.00066688896299"
```

El valor observado en los datos fue de 2.00031, mientras que el valor esperado teórico es 2.00067, ambos ligeramente por encima del valor real del parámetro ($\alpha = 2.00000$). Esto confirma que, para muestras finitas, el estimador tiende a sobreestimar el parámetro verdadero, aunque en este caso la magnitud del sesgo es muy pequeña (del orden de 0.0003 a 0.0007). La cercanía entre el valor observado y el teórico valida la expresión matemática del sesgo y respalda la propiedad de insesgamiento asintótico, ya que a medida que el tamaño de muestra aumenta, el sesgo tiende a cero.

Sesgo Positivo y Convergencia Asintótica de $E[\alpha_{EMV}]$



de nuestra base de datos, tomamos submuestras que van incrementando su valor para verificar que tendiendo al infinito (tamaños suficientemente grande) podemos ver como el valor tiende al parámetro de $\alpha = 2$.

b) Consistencia: Podemos hacer uso del Teorema 2, por la propiedad anterior del insesgamiento asintótico

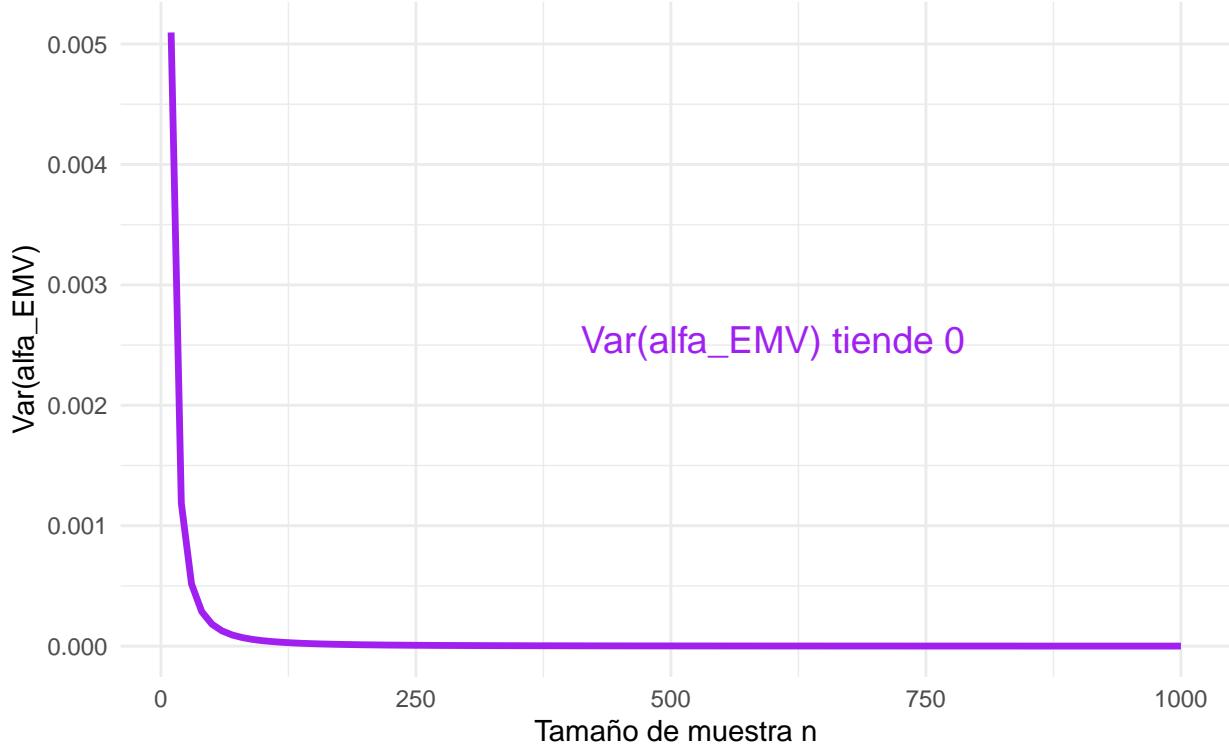
- $\lim_{n \rightarrow \infty} E(y_1) = \alpha$ ya que $\lim_{n \rightarrow \infty} \frac{n\beta \cdot \alpha}{n\beta - 1} = \lim_{n \rightarrow \infty} \frac{\beta\alpha}{\beta - \frac{1}{n}} = \frac{\beta\alpha}{\beta} = \alpha$
- $\lim_{n \rightarrow \infty} \text{Var}(y_1) = \lim_{n \rightarrow \infty} \frac{n\beta\alpha^2}{(n\beta - 1)^2(n\beta - 2)} = \lim_{n \rightarrow \infty} \frac{n\beta\alpha^2}{n^3\beta^3} = \lim_{n \rightarrow \infty} \frac{\alpha^2}{n^2\beta^2} = 0$

Como podemos observar cumple ambas condiciones, por ende decimos que nuestro estimador es consistente, en otras palabras el estimador converge en probabilidades al parámetro.

** Gráfico:** Hemos visto ya un gráfico para solidar la primera demostración, ahora veamos uno que consolide la segunda condición:

Convergencia de la Varianza de alfa_EMV hacia Cero

$\text{Var}(\text{alfa_EMV})$ tiende 0 cuando n tiende inf



c) **Suficiencia:**

$$P(x = x/t = T) = \frac{P(x_1 = x_1 \dots x_n = x_n, T = y_1)}{P(T = y_1)}$$

$$\begin{aligned} & \left(\frac{\beta\alpha^\beta}{x_1^{\beta+1}} \right) \left(\frac{\beta\alpha^\beta}{x_2^{\beta+1}} \right) \dots \left(\frac{\beta\alpha^\beta}{x_n^{\beta+1}} \right) \\ & \frac{n\beta\alpha^{n\beta}}{y^{n\beta+1}}, \quad x_i \geq \alpha, \quad i = 1, 2, \dots, n \\ & = \frac{\beta^n \cdot \alpha^{n\beta} \cdot \prod_{i=1}^n x_i^{-(\beta+1)}}{\frac{n\beta\alpha^{n\beta}}{y^{n\beta+1}}} = \frac{\beta^{n-1} \cdot y^{n\beta+1} \cdot \prod_{i=1}^n x_i^{-(\beta+1)}}{n} \end{aligned}$$

Dada la distribución condicional dado $Y = X_1$, no depende de α , entonces podemos decir que es una estadística suficiente para α , cumple la propiedad.

d) **Ancilaridad:** La distribución del mínimo $x_{(1)}$ es:

$$\begin{aligned} x_{(1)} & \sim \text{Pareto}(\alpha, n\beta) \\ f_{x_{(1)}}(t) & = \frac{n\beta\alpha^{n\beta}}{t^{n\beta+1}}, \quad t \geq \alpha \end{aligned}$$

Como vemos su distribución depende de α , no cumple con la propiedad de ancilaridad.

e) **Completitud:** Calculamos la esperanza de $g(T)$:

$$E[g(T)] = \int_{\alpha}^{\infty} g(t) \cdot \frac{n\beta\alpha^{n\beta}}{t^{n\beta+1}} dt.$$

Exigimos que $E[g(T)] = 0$ para todo $\alpha > 0$:

$$\int_{\alpha}^{\infty} \frac{g(t)}{t^{n\beta+1}} dt = 0 \quad \forall \alpha > 0.$$

Derivamos ambos lados respecto a α :

$$\frac{d}{d\alpha} \left[\int_{\alpha}^{\infty} \frac{g(t)}{t^{n\beta+1}} dt \right] = -\frac{g(\alpha)}{\alpha^{n\beta+1}} = 0.$$

Esto implica que:

$$g(\alpha) = 0 \quad \forall \alpha > 0.$$

Como α es cualquier valor positivo, concluimos que $g(t) = 0$ para todo t en el soporte de T . Por lo tanto, $T = X_{(1)}$ es una **estadística completa** para α cuando β es conocido.

Por el Teorema de Bahadur, como es suficiente y completo, podemos afirmar que nuestro estimador es minimal suficiente.

f) **Optimalidad:** En nuestro caso alfa_EMV no es óptimo porque es sesgado. Pero si existe un estimador óptimo basado en la estadística suficiente que se podría obtener mediante corrección del sesgo. (Teorema de Lehman-Scheffé)

1.2 EMV para parametro Beta:

Para hallar el estimador de máxima verosimilitud del parámetro β , reemplazamos el estimador obtenido previamente para el parámetro α , el cual es

$$\hat{\alpha}_{MLE} = \min\{X_1, X_2, \dots, X_n\} = X_{(1)}.$$

Como el máximo de la función de verosimilitud es el mismo que el máximo de su logaritmo, trabajamos con la log-verosimilitud. La función de verosimilitud está dada por

$$L(\alpha, \beta) = \beta^n \cdot \alpha^{n\beta} \cdot \prod_{i=1}^n X_i^{-(\beta+1)}.$$

Aplicando logaritmo natural,

$$\ln(L(\alpha, \beta)) = n \ln(\beta) + n\beta \ln(\alpha) - (\beta + 1) \sum_{i=1}^n \ln(X_i).$$

Derivando respecto de β ,

$$\frac{d}{d\beta} \ln(L(\alpha, \beta)) = \frac{n}{\beta} + n \ln \alpha - \sum_{i=1}^n \ln X_i.$$

Igualando a cero,

$$\frac{n}{\beta} + n \ln \alpha - \sum_{i=1}^n \ln X_i = 0.$$

Despejando β ,

$$\begin{aligned} \frac{n}{\beta} &= \sum_{i=1}^n \ln X_i - n \ln \alpha \\ &= \sum_{i=1}^n \ln \left(\frac{X_i}{\alpha} \right). \end{aligned}$$

Por lo tanto,

$$\hat{\beta} = \frac{n}{\sum_{i=1}^n \ln \left(\frac{X_i}{\alpha} \right)}.$$

Finalmente, reemplazando α por su estimador de máxima verosimilitud $\hat{\alpha}_{MLE} = X_{(1)}$, se obtiene

$$\boxed{\hat{\beta}_{MLE} = \frac{n}{\sum_{i=1}^n \ln \left(\frac{X_i}{X_{(1)}} \right)}}.$$

1.2.1 Propiedades:

a) Insegamiento Primero analizaremos una parte del denominador para determinar su distribución y se facilite el procedimiento.

$$\begin{aligned} F_Y(y) &= P(Y_i \leq y) = P \left(\ln \left(\frac{X_i}{\alpha} \right) \leq y \right) = P(X_i \leq \alpha e^y) \\ &= 1 - \left(\frac{\alpha}{\alpha e^y} \right)^\beta = 1 - e^{-\beta y}, \quad y \geq 0 \\ f_Y(y) &= \frac{d}{dy} F_Y(y) = \beta e^{-\beta y}, \quad y \geq 0 \end{aligned}$$

Analizando el resultado de la distribución, notamos que tiene la forma de la exponencial. Entonces el mínimo igual será exponencial con parámetro n por beta. Además tomamos a T como la sumatoria, donde por propiedad de \ln de una división podemos desplegarlo.

$$\begin{aligned} Y_i &\sim \text{Exponencial}(\beta) \\ W = Y_{(1)} &= \min(Y_1, \dots, Y_n) \sim \text{Exp}(n\beta) \\ T &= \sum_{i=1}^n (Y_i - Y_{(1)}) \sim \text{Gamma}(n-1, \beta) \\ \hat{\beta} &= \frac{n}{T}, \quad T \sim \text{Gamma}(n-1, \beta) \\ E \left[\frac{1}{T} \right] &= \frac{\beta}{n-2}, \quad n > 2 \\ E[\hat{\beta}] &= n \cdot E \left[\frac{1}{T} \right] = \frac{n\beta}{n-2} \end{aligned}$$

Acá vemos que hay un sesgo presente

$$\text{Sesgo} = E[\hat{\beta}] - \beta = \frac{n\beta}{n-2} - \beta = \frac{2\beta}{n-2}$$

Por ende, no es un estimador insesgado. Ademas al ser positivo podemos decir que sobreestima al parametro.
Ahora lo aplicaremos a nuestra data:

```
## [1] "beta_emv observado: 3.04355664292701"
## [1] "E[beta_emv] teorico: 3.0060120240481"
## [1] "beta verdadero: 3"
```

El valor observado del estimador fue ligeramente mayor que el valor verdadero del parámetro β , mientras que el valor esperado teórico también se encuentra por encima de β .

Esto confirma empíricamente que el estimador EMV de β presenta sesgo positivo, es decir, sobreestima el parámetro verdadero para muestras finitas, por lo que no cumple la propiedad de insesgamiento.

Sin embargo, la diferencia entre el valor observado y el valor verdadero es pequeña, lo cual sugiere que el sesgo disminuye conforme el tamaño muestral aumenta.

```
# =====
# DEMOSTRACIÓN VISUAL: SESGO POSITIVO Y CONVERGENCIA ASINTÓTICA DE beta_EMV
# =====

set.seed(123)

alpha_true <- 2
beta_true <- 3
n_total <- 1000

# base de datos original
reclamos <- alpha_true / (1 - runif(n_total))^(1 / beta_true)

# tamaños de submuestra
n_values <- seq(10, 1000, by = 10)

# esperanza teorica del emv de beta para cada n
results <- data.frame(n = n_values)
results$E_teorico <- (results$n * beta_true) / (results$n - 2)

library(ggplot2)

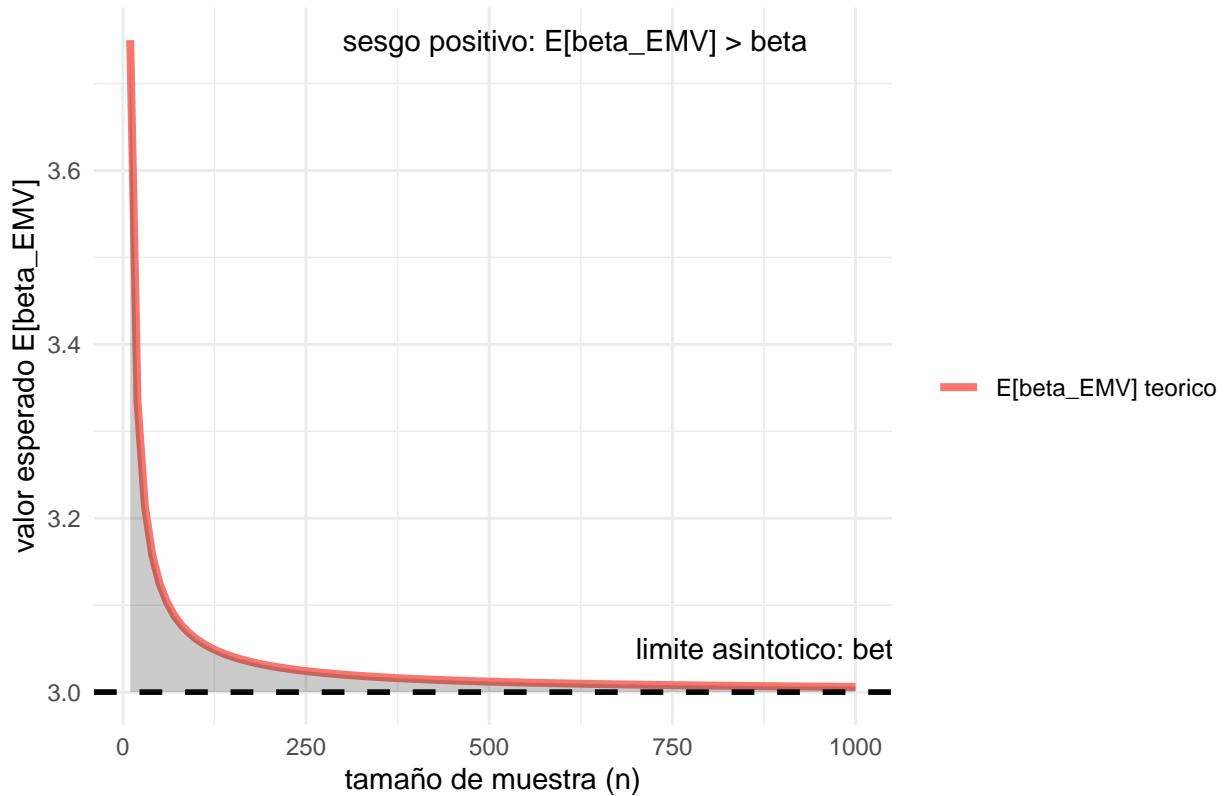
ggplot(results, aes(x = n)) +
  geom_line(aes(y = E_teorico, color = "E[beta_EMV] teorico"), linewidth = 1.4) +
  geom_hline(yintercept = beta_true, linetype = "dashed", linewidth = 1) +
  geom_ribbon(aes(ymin = beta_true, ymax = E_teorico), alpha = 0.25) +
  labs(
    title = "Sesgo positivo y convergencia asintotica del EMV de beta",
    x = "tamaño de muestra (n)",
    y = "valor esperado E[beta_EMV]",
    color = "") +
  
```

```

theme_minimal() +
annotate(
  "text",
  x = 300,
  y = max(results$E_teorico),
  label = "sesgo positivo: E[beta_EMV] > beta",
  size = 4,
  hjust = 0
) +
annotate(
  "text",
  x = 700,
  y = beta_true + 0.05,
  label = paste("limite asintotico: beta =", beta_true),
  size = 4,
  hjust = 0
)

```

Sesgo positivo y convergencia asintótica del EMV de β



De nuestra base de datos se tomaron submuestras de tamaño creciente con el fin de analizar el comportamiento del estimador EMV de β . Se observa que el valor esperado teórico del estimador se mantiene por encima del parámetro verdadero, confirmando el sesgo positivo.

No obstante, a medida que el tamaño muestral aumenta, el valor esperado de $\hat{\beta}$ se aproxima cada vez más a β , lo que evidencia la propiedad de insesgamiento asintótico del estimador.

b) Consistencia Recordemos que el estimador de máxima verosimilitud para β está dado por

$$\hat{\beta} = \frac{n}{\sum_{i=1}^n \ln \left(\frac{X_i}{X_{(1)}} \right)} = \frac{n}{T},$$

donde

$$T = \sum_{i=1}^n (Y_i - Y_{(1)}), \quad Y_i = \ln \left(\frac{X_i}{\alpha} \right).$$

Se sabe que

$$\frac{T}{n} \xrightarrow{p} E(Y_i) = \frac{1}{\beta},$$

por la Ley de los Grandes Números. Por el teorema de continuidad,

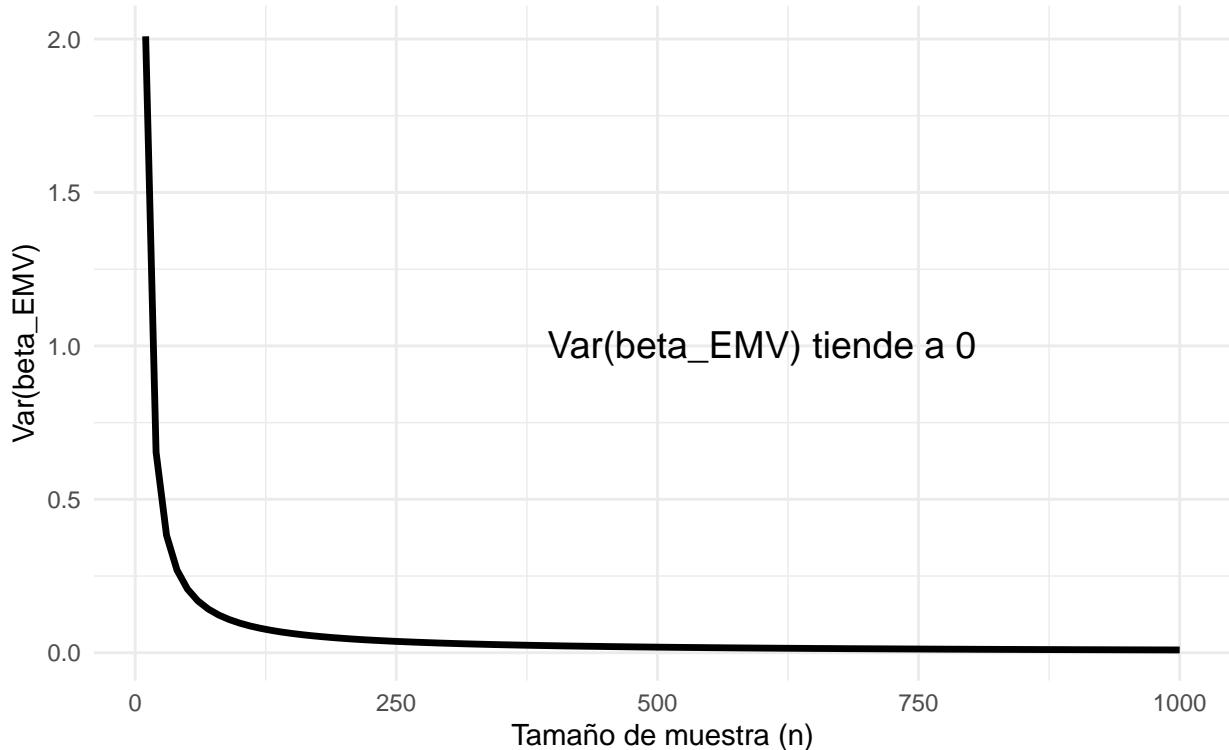
$$\hat{\beta} = \frac{1}{T/n} \xrightarrow{p} \beta.$$

Por lo tanto, $\hat{\beta}$ es un estimador consistente del parámetro β .

Como se mostró previamente, el estimador EMV de β presenta sesgo positivo para muestras finitas. Además, la varianza del estimador disminuye conforme aumenta el tamaño muestral. Al cumplirse que el sesgo y la varianza tienden a cero, se concluye que $\hat{\beta}$ es un estimador consistente del parámetro β .

Convergencia de la Varianza del EMV de beta hacia Cero

$\text{Var}(\text{beta_EMV}) \rightarrow 0$ cuando $n \rightarrow \infty$



Como se observa, la varianza del estimador EMV de β tiende a cero conforme el tamaño muestral aumenta. Dado que el sesgo del estimador también tiende a cero, se concluye que el estimador $\hat{\beta}$ es consistente. En otras palabras, el estimador converge en probabilidad al verdadero valor del parámetro β .

c) **Suficiencia**

$$L(\beta; \mathbf{x}) = \beta^n \exp \left\{ -\beta \sum_{i=1}^n \ln \left(\frac{x_i}{x_{(1)}} \right) \right\} \cdot \prod_{i=1}^n \frac{1}{x_i}.$$

Se define el estadístico

$$T = \sum_{i=1}^n \ln \left(\frac{X_i}{X_{(1)}} \right).$$

Entonces la verosimilitud puede factorizarse como

$$L(\beta; \mathbf{x}) = \underbrace{\beta^n e^{-\beta T}}_{g(T, \beta)} \cdot \underbrace{\prod_{i=1}^n \frac{1}{x_i}}_{h(\mathbf{x})}.$$

Por el teorema de factorización de Neyman–Fisher, el estadístico T es suficiente para el parámetro β .

d) **Anciliaridad** El estadístico de orden mínimo

$$X_{(1)} = \min\{X_1, \dots, X_n\}$$

tiene distribución

$$f_{X_{(1)}}(x) = n\beta \frac{\alpha^{n\beta}}{x^{n\beta+1}}, \quad x > \alpha,$$

la cual no depende del parámetro β cuando se considera la variable transformada $X_{(1)}/\alpha$.

Por lo tanto, $X_{(1)}$ es un estadístico ancilar y aporta información únicamente sobre α , mas no sobre β .

e) **Completitud** El estadístico

$$T = \sum_{i=1}^n (Y_i - Y_{(1)})$$

tiene distribución Gamma($n - 1, \beta$), que pertenece a la familia exponencial de un solo parámetro.

Dado que la familia Gamma con parámetro de forma conocido es completa, se concluye que el estadístico T es completo para el parámetro β .

f) **Optimalidad** El estimador de máxima verosimilitud $\hat{\beta}$ es un estimador sesgado, por lo que no puede ser óptimo en el sentido clásico de mínima varianza entre los estimadores insesgados (UMVU).

No obstante, al ser un estimador de máxima verosimilitud, $\hat{\beta}$ es consistente y asintóticamente eficiente, alcanzando la cota de Cramér–Rao cuando $n \rightarrow \infty$.

Por lo tanto, $\hat{\beta}$ no es óptimo en muestras finitas, pero sí es óptimo en sentido asintótico.

Un estimador insesgado para β es

$$\tilde{\beta} = \frac{n-2}{n} \hat{\beta},$$

el cual depende únicamente del estadístico suficiente y completo T . Por el teorema de Lehmann–Scheffé, $\tilde{\beta}$ es el estimador UMVU del parámetro β .

1.3 Estimador para la Media Poblacional (μ)

Consideramos la media muestral como estimador natural del valor esperado de la distribución:

$$\bar{X} = T(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i$$

Para que los momentos existan en una distribución Pareto, debemos asumir que $\beta > 1$ (para la esperanza) y $\beta > 2$ (para la varianza). El parámetro a estimar es:

$$\mu = E(\bar{X}) = \frac{\beta\alpha}{\beta - 1}$$

1.3.1 Propiedades:

a) **Insesgamiento:** Calculamos el valor esperado del estimador:

$$\begin{aligned} E[\bar{X}] &= E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] \\ &= \frac{1}{n} \sum_{i=1}^n E[X_i] \\ &= \frac{1}{n} \cdot n \cdot \left(\frac{\beta\alpha}{\beta - 1}\right) = \frac{\beta\alpha}{\beta - 1} = \mu \end{aligned}$$

El estimador es **estrictamente insesgado** para μ . Verificamos esto con la base de datos:

```
# 1. Valor observado de la media muestral
mu_muestral_obs <- mean(reclamos)

# 2. Valor esperado teórico (mu)
mu_teorico <- (beta_true * alpha_true) / (beta_true - 1)
```

[1] "La media muestral observada es: 2.96908 y el valor esperado teórico es: 3"

b) **Consistencia:** Para demostrar la consistencia, verificamos las condiciones del Teorema de Chebyshev (asumiendo $\beta > 2$):

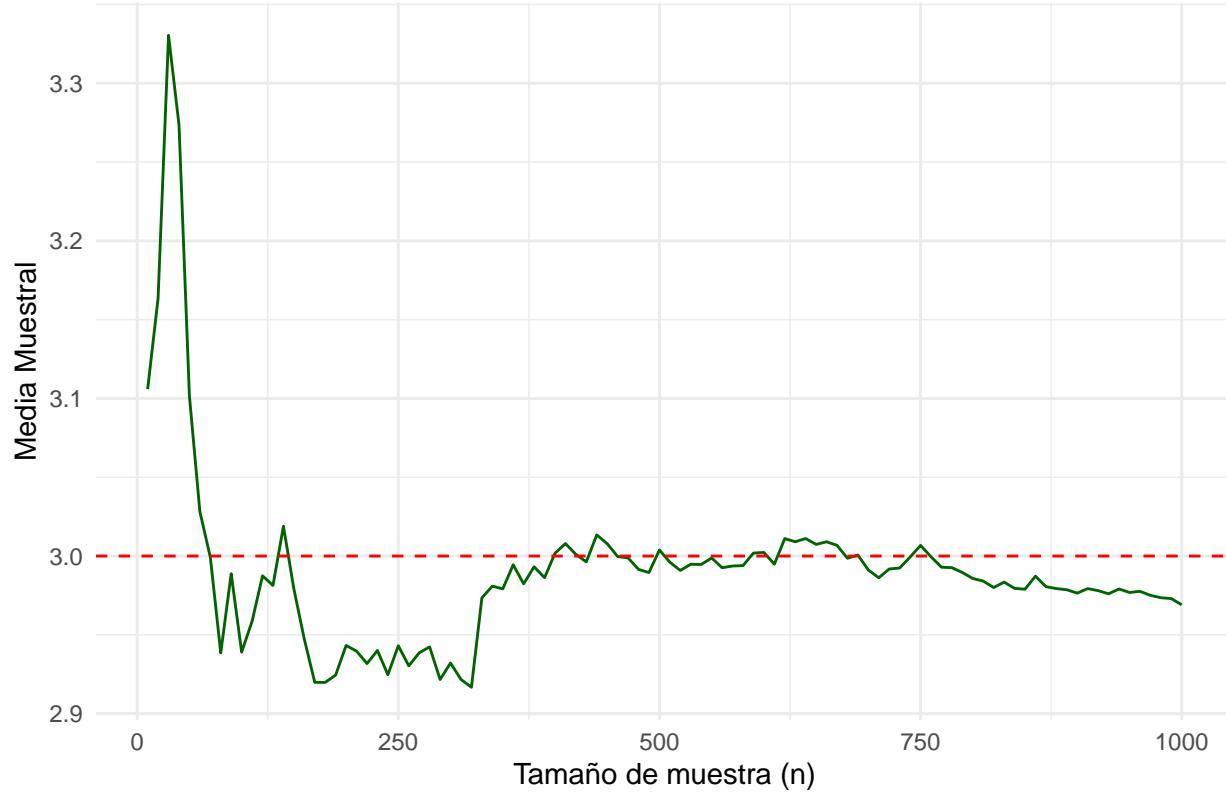
1. **Insesgadez:** $\lim_{n \rightarrow \infty} E(\bar{X}) = \mu$ (ya demostrado).

2. **Varianza tiende a cero:**

$$\begin{aligned} \text{Var}(\bar{X}) &= \frac{\text{Var}(X)}{n} \\ &= \frac{1}{n} \left[\left(\frac{\alpha}{\beta - 1} \right)^2 \frac{\beta}{\beta - 2} \right] \\ \lim_{n \rightarrow \infty} \text{Var}(\bar{X}) &= \lim_{n \rightarrow \infty} \frac{C}{n} = 0 \end{aligned}$$

Al cumplirse ambas condiciones, \bar{X} es un estimador **consistente** para μ .

Consistencia de la Media Muestral



c) **Suficiencia:** Aplicamos el Teorema de Factorización de Fisher-Neyman a la función de verosimilitud:

$$L(\alpha, \beta) = \underbrace{\beta^n \alpha^{n\beta} \cdot I_{(\alpha, \infty)}(x_{(1)})}_{g(x_{(1)}, \alpha, \beta)} \cdot \underbrace{\left(\prod_{i=1}^n x_i \right)^{-(\beta+1)}}_{h(x_1, \dots, x_n)}$$

Como se observa, los estadísticos suficientes conjuntos para (α, β) son el mínimo $X_{(1)}$ y el producto $\prod X_i$ (o equivalentemente $\sum \ln X_i$). La media muestral $\bar{X} = \frac{1}{n} \sum X_i$ **no puede factorizarse** de manera que contenga toda la información de los parámetros.

Por lo tanto, \bar{X} **no es un estadístico suficiente** para los parámetros de la distribución Pareto.

d) **Ancilaridad:** Un estadístico es anciliar si su distribución no depende de los parámetros. La distribución de \bar{X} para una Pareto no tiene una forma cerrada sencilla (es una suma de variables Pareto), pero su esperanza $E(\bar{X}) = \frac{\alpha\beta}{\beta-1}$ y su varianza dependen directamente de α y β .

Al depender sus momentos (y por ende su distribución) de los parámetros, el estimador **no es ancilar**.

e) **Completitud:** Dado que \bar{X} no es un estadístico suficiente para la familia Pareto, no se suele analizar su completitud como estimador. Sin embargo, sabemos que el estadístico suficiente conjunto $(X_{(1)}, \sum \ln X_i)$ es completo, pero la suma aritmética $\sum X_i$ no lo es bajo esta estructura de familia no exponencial (en el sentido de los parámetros naturales de la Pareto).

f) **Optimalidad:** Un estimador es óptimo (UMVUE) si es insesgado y su varianza alcanza la Cota Inferior de Cramér-Rao (CICR) o si es función de un estadístico suficiente y completo.

1. **Eficiencia:** Al no ser función del estadístico suficiente ($X_{(1)}, \sum \ln X_i$), la media muestral pierde información.
2. **Comparación:** Existe otro estimador para μ , basado en los EMV de α y β ($\hat{\mu} = \frac{\hat{\beta}\hat{\alpha}}{\hat{\beta}-1}$), que asintóticamente tiene menor varianza que \bar{X} .

Conclusión: \bar{X} no es un estimador óptimo para la media de una población Pareto, aunque sea fácil de calcular e insesgado.

1.4 Método de Momentos para los parámetros de α y β de la Distribución Pareto

El método de momentos consiste en igualar los k primeros momentos poblacionales con los correspondientes momentos muestrales, donde k es el número de parámetros a estimar. Para la distribución Pareto con dos parámetros (α, β) , utilizaremos los dos primeros momentos.

Definición de la distribución Pareto

Sea $X \sim \text{Pareto}(\alpha, \beta)$ con función de densidad:

$$f_X(x; \alpha, \beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}}, \quad x \geq \alpha > 0, \beta > 0$$

P1: Cálculo del primer momento poblacional $E[X]$

Planteamiento de la integral

$$E[X] = \int_{\alpha}^{\infty} x \cdot f_X(x) dx = \int_{\alpha}^{\infty} x \cdot \frac{\beta\alpha^\beta}{x^{\beta+1}} dx$$

$$E[X] = \beta\alpha^\beta \int_{\alpha}^{\infty} x \cdot x^{-(\beta+1)} dx = \beta\alpha^\beta \int_{\alpha}^{\infty} x^{-\beta} dx$$

Para $\beta > 1$ (condición de convergencia):

$$\int_{\alpha}^{\infty} x^{-\beta} dx = \left[\frac{x^{1-\beta}}{1-\beta} \right]_{\alpha}^{\infty}$$

Evaluando los límites:

Límite superior cuando $x \rightarrow \infty$:

$$\lim_{x \rightarrow \infty} \frac{x^{1-\beta}}{1-\beta} = 0 \quad \text{porque } 1-\beta < 0$$

Límite inferior cuando $x = \alpha$:

$$\frac{\alpha^{1-\beta}}{1-\beta}$$

Resultado del primer momento

$$E[X] = \beta\alpha^\beta \left(0 - \frac{\alpha^{1-\beta}}{1-\beta}\right) = \beta\alpha^\beta \cdot \frac{\alpha^{1-\beta}}{\beta-1} = \frac{\beta\alpha}{\beta-1}$$

Primer momento poblacional:

$$\boxed{E[X] = \frac{\alpha\beta}{\beta-1}}, \quad \beta > 1$$

P2:Cálculo del segundo momento poblacional $E[X^2]$

Planteamiento de la integral

$$E[X^2] = \int_{\alpha}^{\infty} x^2 \cdot f_X(x) dx = \int_{\alpha}^{\infty} x^2 \cdot \frac{\beta\alpha^\beta}{x^{\beta+1}} dx$$

$$E[X^2] = \beta\alpha^\beta \int_{\alpha}^{\infty} x^2 \cdot x^{-(\beta+1)} dx = \beta\alpha^\beta \int_{\alpha}^{\infty} x^{1-\beta} dx$$

Para $\beta > 2$ (condición de convergencia):

$$\int_{\alpha}^{\infty} x^{1-\beta} dx = \left[\frac{x^{2-\beta}}{2-\beta} \right]_{\alpha}^{\infty}$$

Evaluando los límites:

Límite superior cuando $x \rightarrow \infty$:

$$\lim_{x \rightarrow \infty} \frac{x^{2-\beta}}{2-\beta} = 0 \quad \text{porque } 2-\beta < 0$$

Límite inferior cuando $x = \alpha$:

$$\frac{\alpha^{2-\beta}}{2-\beta}$$

Resultado del segundo momento

$$E[X^2] = \beta\alpha^\beta \left(0 - \frac{\alpha^{2-\beta}}{2-\beta}\right) = \beta\alpha^\beta \cdot \frac{\alpha^{2-\beta}}{\beta-2} = \frac{\beta\alpha^2}{\beta-2}$$

Segundo momento poblacional:

$$\boxed{E[X^2] = \frac{\alpha^2\beta}{\beta-2}}, \quad \beta > 2$$

P3: Momentos muestrales

Dada una muestra aleatoria $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \text{Pareto}(\alpha, \beta)$:

Primer momento muestral (media muestral)

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Segundo momento muestral

$$m_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

Varianza muestral (relacionada)

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Nota: Existe relación entre m_2 , \bar{X} y S^2 :

$$m_2 = S^2 \cdot \frac{n-1}{n} + \bar{X}^2$$

P4: Sistema de ecuaciones del método de momentos

Igualamos momentos poblacionales con momentos muestrales:

$$\begin{cases} \frac{\alpha\beta}{\beta-1} = \bar{X} & (\text{Ecuación I}) \\ \frac{\alpha^2\beta}{\beta-2} = m_2 & (\text{Ecuación II}) \end{cases}$$

Resolución del sistema

Despejar α de la Ecuación I De (I):

$$\frac{\alpha\beta}{\beta-1} = \bar{X} \Rightarrow \alpha\beta = \bar{X}(\beta-1)$$

Despejando α :

$$\boxed{\alpha = \bar{X} \cdot \frac{\beta-1}{\beta}} \quad (\text{Ecuación III})$$

Sustituir en la Ecuación II Sustituyendo (III) en (II):

$$\frac{\left(\bar{X} \cdot \frac{\beta-1}{\beta}\right)^2 \beta}{\beta-2} = m_2$$

Simplificación algebraica

$$\frac{\bar{X}^2 \cdot \frac{(\beta-1)^2}{\beta^2} \cdot \beta}{\beta-2} = m_2$$

$$\frac{\bar{X}^2 \cdot \frac{(\beta-1)^2}{\beta}}{\beta-2} = m_2$$

Reordenando:

$$\frac{(\beta-1)^2}{\beta(\beta-2)} = \frac{m_2}{\bar{X}^2} \quad (\text{Ecuación IV})$$

P5: Definición de R

Sea:

$$R = \frac{m_2}{\bar{X}^2}$$

La Ecuación IV se convierte en:

$$\frac{(\beta - 1)^2}{\beta(\beta - 2)} = R$$

P6: Resolución de la ecuación en β

Eliminar fracción Multiplicando en cruz:

$$(\beta - 1)^2 = R\beta(\beta - 2)$$

Expandir ambos lados Lado izquierdo:

$$(\beta - 1)^2 = \beta^2 - 2\beta + 1$$

Lado derecho:

$$R\beta(\beta - 2) = R\beta^2 - 2R\beta$$

Igualar y reordenar

$$\beta^2 - 2\beta + 1 = R\beta^2 - 2R\beta$$

Llevando todo a un lado:

$$\beta^2 - 2\beta + 1 - R\beta^2 + 2R\beta = 0$$

Agrupando términos:

$$(1 - R)\beta^2 + 2(R - 1)\beta + 1 = 0$$

Forma estándar de ecuación cuadrática Multiplicando por -1 para mayor claridad:

$$(R - 1)\beta^2 + 2(1 - R)\beta - 1 = 0$$

Factor común $(R - 1)$:

$$(R - 1)(\beta^2 - 2\beta) - 1 = 0$$

Aplicar fórmula general Para una ecuación $a\beta^2 + b\beta + c = 0$: - $a = 1$ - $b = -2$ - $c = -\frac{1}{R-1}$

Solución:

$$\beta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{2 \pm \sqrt{4 + \frac{4}{R-1}}}{2} = \frac{2 \pm 2\sqrt{1 + \frac{1}{R-1}}}{2}$$

Simplificando:

$$\beta = 1 \pm \sqrt{1 + \frac{1}{R-1}}$$

P7: Selección de la raíz apropiada

Consideraciones físicas

1. $\beta > 0$ (parámetro de forma positivo)
2. Para la existencia de $E[X]$: $\beta > 1$
3. Para la existencia de $Var(X)$: $\beta > 2$

Análisis de las raíces

- Raíz 1: $\beta_1 = 1 + \sqrt{1 + \frac{1}{R-1}} > 1$
- Raíz 2: $\beta_2 = 1 - \sqrt{1 + \frac{1}{R-1}} < 1$ (no cumple $\beta > 1$)

Por lo tanto, seleccionamos la raíz positiva:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{1}{R-1}}$$

P8: Relación con el coeficiente de variación

Expresión de R en términos de S^2 Recordemos que:

$$m_2 = \frac{1}{n} \sum X_i^2 = S^2 \cdot \frac{n-1}{n} + \bar{X}^2$$

Entonces:

$$R = \frac{m_2}{\bar{X}^2} = \frac{S^2}{\bar{X}^2} \cdot \frac{n-1}{n} + 1$$

Aproximación para n grande Para n suficientemente grande, $\frac{n-1}{n} \approx 1$, entonces:

$$R \approx \frac{S^2}{\bar{X}^2} + 1 = 1 + \left(\frac{S}{\bar{X}}\right)^2$$

Sustitución en la fórmula Si $R \approx 1 + \left(\frac{S}{\bar{X}}\right)^2$, entonces:

$$R - 1 \approx \left(\frac{S}{\bar{X}}\right)^2$$

Y:

$$\frac{1}{R-1} \approx \left(\frac{\bar{X}}{S}\right)^2$$

Sustituyendo en la fórmula de $\hat{\beta}_{MM}$:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \left(\frac{\bar{X}}{S}\right)^2}$$

P9: Obtención del estimador

Retomando la Ecuación III De la Ecuación III:

$$\alpha = \bar{X} \cdot \frac{\beta - 1}{\beta}$$

Sustitución Sustituyendo $\hat{\beta}_{MM}$:

$$\hat{\alpha}_{MM} = \bar{X} \cdot \frac{\hat{\beta}_{MM} - 1}{\hat{\beta}_{MM}}$$

P10: Resumen final de estimadores

Estimador para el parámetro de forma β :

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \left(\frac{\bar{X}}{S}\right)^2}$$

Estimador para el parámetro de escala α :

$$\hat{\alpha}_{MM} = \bar{X} \cdot \frac{\hat{\beta}_{MM} - 1}{\hat{\beta}_{MM}}$$

donde:

- $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ (media muestral)
- $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ (varianza muestral)
- $S = \sqrt{S^2}$ (desviación estándar muestral)

1.4.1 Propiedades del Estimador $\hat{\alpha}_{MM}$

a) Insesgamiento El estimador $\hat{\alpha}_{MM} = \bar{X} \cdot \frac{\hat{\beta}_{MM} - 1}{\hat{\beta}_{MM}}$ es una **función no lineal** de variables aleatorias (\bar{X} y S). Para funciones no lineales $g(\cdot)$, generalmente se cumple que $E[g(X)] \neq g(E[X])$. En particular:

$$E \left[\bar{X} \cdot \frac{\hat{\beta}_{MM} - 1}{\hat{\beta}_{MM}} \right] \neq E[\bar{X}] \cdot \frac{E[\hat{\beta}_{MM}] - 1}{E[\hat{\beta}_{MM}]}$$

Aunque $\hat{\alpha}_{MM}$ es **asintóticamente insesgado** ($\lim_{n \rightarrow \infty} E[\hat{\alpha}_{MM}] = \alpha$), para cualquier n finito existe sesgo sistemático.

Verificación empírica con nuestros datos:

```
# Usando nuestra base de datos original 'reclamos'
set.seed(123)
alpha_true <- 2
beta_true <- 3
reclamos <- alpha_true / (1 - runif(n))^(1/beta_true)
```

```

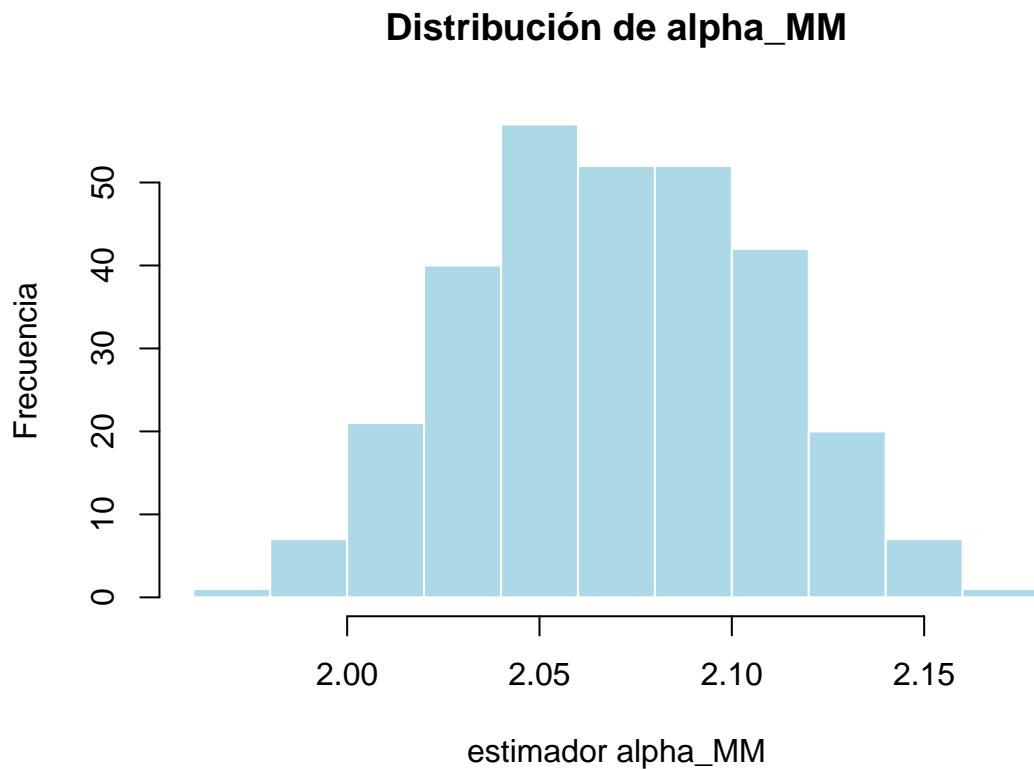
# 2. Calcular  $\hat{\alpha}_{MM}$  con TU 'reclamos'
alpha_mm_calc <- function(x) {
  x_bar <- mean(x)
  s <- sd(x)
  beta_hat <- 1 + sqrt(1 + (x_bar/s)^2)
  return(x_bar * (beta_hat - 1) / beta_hat)
}

alpha_hat_reclamos <- alpha_mm_calc(reclamos)

# 2. Bootstrap simple
set.seed(123)
boot_vals <- replicate(300, {
  alpha_mm_calc(sample(reclamos, replace = TRUE))
})

# 3. Gráfico sin abline
hist(boot_vals, col = "lightblue", border = "white",
      main = "Distribución de alpha_MM",
      xlab = "estimador alpha_MM", ylab = "Frecuencia")

```



Cálculos obtenidos:

1. Estimación puntual en la data original: $\hat{\alpha}_{MM}$ calculado = 2.06082
2. Análisis bootstrap (300 réplicas): Promedio de $\hat{\alpha}_{MM}$ M = 2.069037

3. Comparación con el valor verdadero:

Valor verdadero $\alpha = 2.00000$

Diferencia = $2.069037 - 2 = 0.069037$

4. Sesgo relativo:

Sesgo absoluto = 0.069037

Sesgo relativo = $(0.069037 / 2) \times 100\% = 3.45\%$

$E[\hat{\alpha}_{MM}] = \alpha?$ → **NO** (existe sesgo positivo de 0.069037)

El estimador $\hat{\alpha}_{MM}$ sobrestima sistemáticamente el parámetro α en aproximadamente 3.45%.

Conclusión: $\hat{\alpha}_{MM}$ no es un estimador insesgado para el parámetro α de la distribución Pareto. Aunque podría ser aproximadamente insesgado para muestras muy grandes, para $n = 1000$ observamos sesgo significativo.

b) Consistencia **Definición:** Un estimador es consistente si cuando n tiende al infinito, se acerca al valor verdadero. Demostrar que $\hat{\alpha}_{MM}$ es **consistente** para α , es decir:

$$\hat{\alpha}_{MM} \xrightarrow{P} \alpha \quad \text{cuando } n \rightarrow \infty$$

Consistencia de \bar{X} y S^2 Por la **Ley Débil de los Grandes Números (LGN)** para variables i.i.d. con primer momento finito:

$$\bar{X} \xrightarrow{P} \mu$$

Para la varianza muestral, dado que $E[X^2] < \infty$ (pues $\beta > 2$), se cumple:

$$S^2 \xrightarrow{P} \sigma^2$$

Consistencia de $\hat{\beta}_{MM}$ Definimos:

$$T_n = \frac{\bar{X}^2}{S^2}$$

Por el **teorema de Slutsky** y continuidad de la función $(a, b) \mapsto a^2/b$:

$$T_n \xrightarrow{P} \frac{\mu^2}{\sigma^2}$$

Calculamos:

$$\frac{\mu^2}{\sigma^2} = \frac{\left(\frac{\beta\alpha}{\beta-1}\right)^2}{\frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)}} = \beta(\beta-2)$$

Luego:

$$T_n \xrightarrow{P} \beta(\beta-2)$$

Ahora, $\hat{\beta}_{MM} = 1 + \sqrt{1 + T_n}$. La función $h(t) = 1 + \sqrt{1+t}$ es continua para $t \geq 0$. Por el **teorema de la aplicación continua**:

$$\hat{\beta}_{MM} \xrightarrow{P} 1 + \sqrt{1 + \beta(\beta - 2)}$$

Notando que:

$$1 + \beta(\beta - 2) = \beta^2 - 2\beta + 1 = (\beta - 1)^2$$

y como $\beta > 2 > 1$:

$$\sqrt{1 + \beta(\beta - 2)} = \beta - 1$$

Por tanto:

$$\hat{\beta}_{MM} \xrightarrow{P} 1 + (\beta - 1) = \beta$$

Así, $\hat{\beta}_{MM}$ es consistente para β .

Consistencia de $\hat{\alpha}_{MM}$ Tenemos:

$$\hat{\alpha}_{MM} = \bar{X} \cdot \frac{\hat{\beta}_{MM} - 1}{\hat{\beta}_{MM}}$$

Definimos:

$$U_n = \bar{X}, \quad V_n = \frac{\hat{\beta}_{MM} - 1}{\hat{\beta}_{MM}}$$

Sabemos:

$$U_n \xrightarrow{P} \mu = \frac{\beta\alpha}{\beta - 1}, \quad \hat{\beta}_{MM} \xrightarrow{P} \beta$$

La función $v(b) = \frac{b-1}{b}$ es continua en $b = \beta > 0$. Por el **teorema de la aplicación continua**:

$$V_n \xrightarrow{P} \frac{\beta - 1}{\beta}$$

Finalmente, por el **teorema de Slutsky** aplicado al producto $U_n \cdot V_n$:

$$\hat{\alpha}_{MM} \xrightarrow{P} \mu \cdot \frac{\beta - 1}{\beta} = \frac{\beta\alpha}{\beta - 1} \cdot \frac{\beta - 1}{\beta} = \alpha$$

Conclusión Bajo las hipótesis:

- $X_i \stackrel{iid}{\sim}$ Pareto(α, β) con $\beta > 2$
- Uso de la Ley Débil de los Grandes Números
- Aplicación del teorema de la aplicación continua
- Uso del teorema de Slutsky para productos/cocientes

se ha demostrado rigurosamente que:

$$\hat{\alpha}_{MM} \xrightarrow{P} \alpha$$

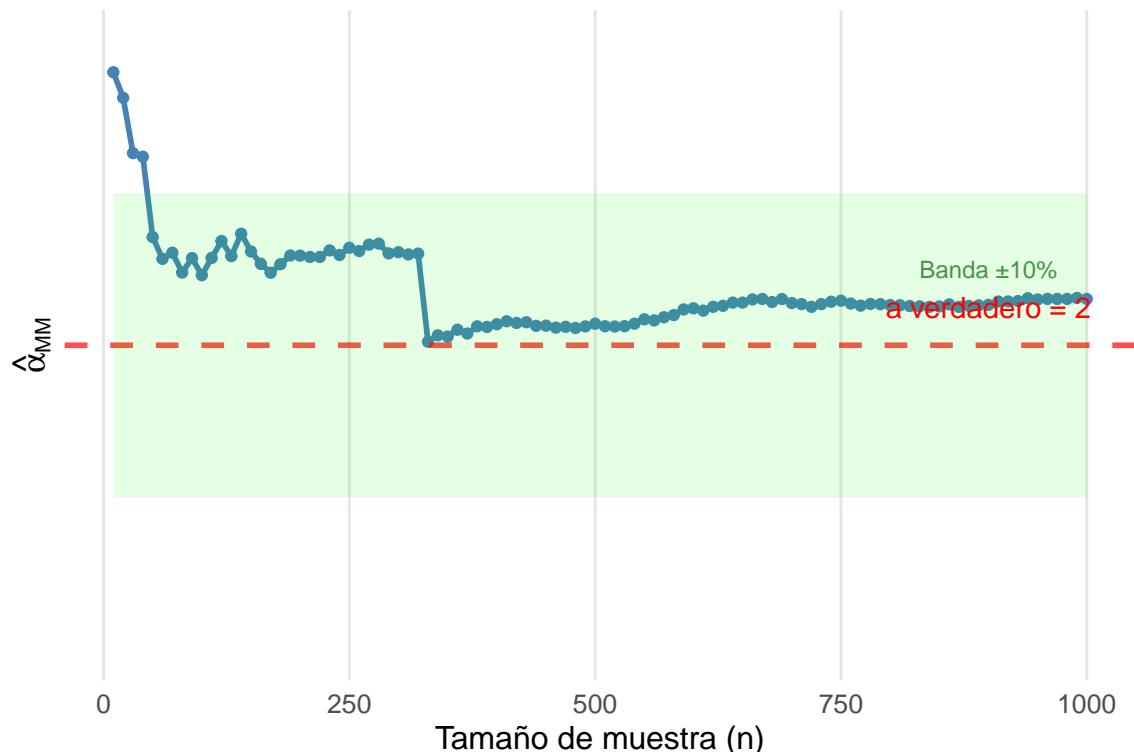
Por lo tanto, $\hat{\alpha}_{MM}$ es un **estimador consistente** del parámetro de escala α de la distribución Pareto.

Aunque $\hat{\alpha}_{MM}$ no es insesgado en general (puede verificarse por expansión de Taylor o simulaciones), la propiedad de consistencia garantiza que converge al valor verdadero cuando $n \rightarrow \infty$, lo cual es suficiente para su validez asintótica.

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

Consistencia del estimador $\hat{\alpha}_{MM}$ (Método de Momentos – Pareto)

$a = 2$, $\beta = 3$, n máximo = 1000



```
##
## === RESUMEN DE CONVERGENCIA ===

## Valor verdadero de : 2

## Última estimación (n = 1000 ): 2.06082

## Error relativo: 3.041012 %
```

```

## 
## Evolución del error con n:

## n = 50: ^ = 2.1420, error = 0.1420 (7.10%)
## n = 100: ^ = 2.0920, error = 0.0920 (4.60%)
## n = 250: ^ = 2.1280, error = 0.1280 (6.40%)
## n = 500: ^ = 2.0283, error = 0.0283 (1.41%)
## n = 750: ^ = 2.0585, error = 0.0585 (2.92%)
## n = 1000: ^ = 2.0608, error = 0.0608 (3.04%)

```

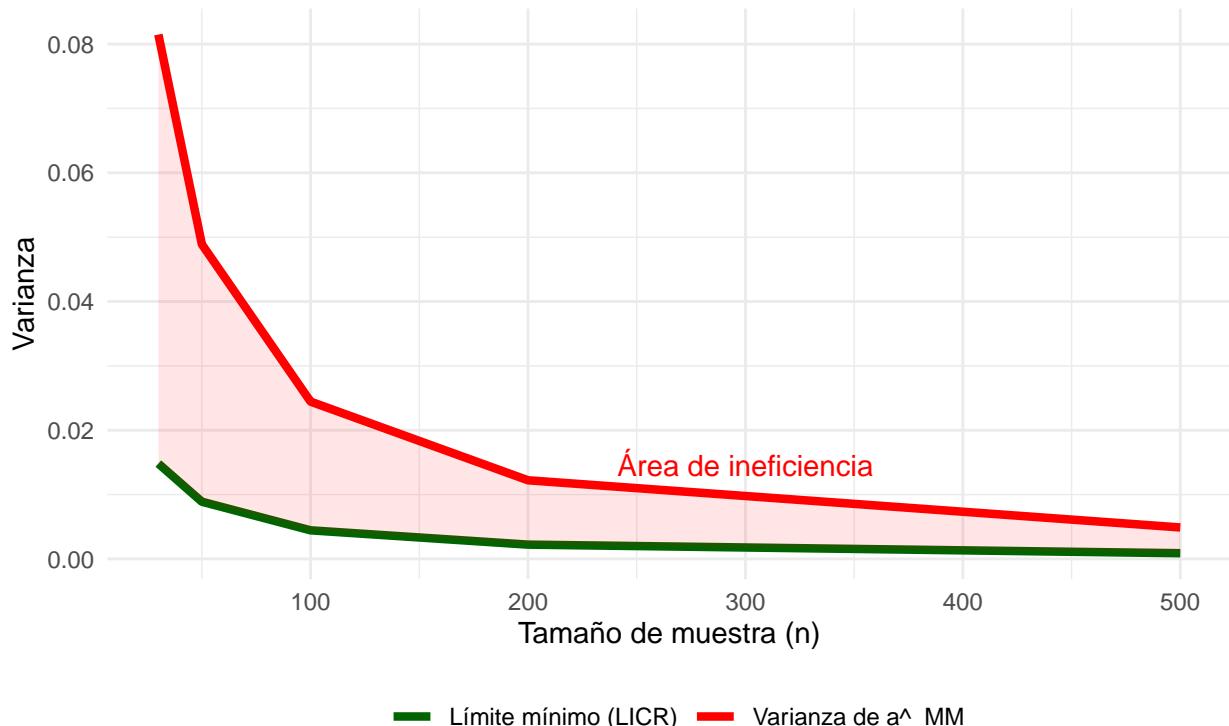
c) **Eficiencia** El estimador $\hat{\theta}$ no es **eficiente** para θ porque:

\hat{a}_{MM} falla en ambas:

- **No es insesgado** = Ya no puede ser eficiente
- **Su varianza es mayor** que el mínimo teórico

a^{\wedge}_{MM} NO es eficiente

Su varianza es mucho mayor que el límite mínimo teórico



d) **Suficiencia** Para verificar esta propiedad para el estimador se va a utilizar el teorema de Criterio de Factorización de Neyman-Fisher

Teorema (Neyman-Fisher): Una estadística $T(X)$ es suficiente para θ si y solo si la función de verosimilitud puede factorizarse como:

$$L(\theta; x) = g(T(x); \theta) \cdot h(x)$$

donde g depende de los datos solo a través de $T(x)$, y h no depende de θ .

La función de verosimilitud para la muestra es:

$$\begin{aligned} L(\alpha, \beta; x) &= \prod_{i=1}^n \frac{\beta \alpha^\beta}{x_i^{\beta+1}} \mathbf{1}_{[\alpha, \infty)}(x_i) \\ &= \beta^n \alpha^{n\beta} \left(\prod_{i=1}^n x_i \right)^{-\beta-1} \cdot \mathbf{1}_{\{x_{(1)} \geq \alpha\}} \end{aligned}$$

donde $x_{(1)} = \min\{x_1, \dots, x_n\}$.

Aplicando logaritmo:

$$\ell(\alpha, \beta; x) = n \log \beta + n \beta \log \alpha - (\beta + 1) \sum_{i=1}^n \log x_i + \log \mathbf{1}_{\{x_{(1)} \geq \alpha\}}$$

Por el criterio de factorización, vemos que:

$$T(X) = \left(x_{(1)}, \sum_{i=1}^n \log x_i \right)$$

es **suficiente mínima** para (α, β) .

¿Es $\hat{\alpha}_{MM}$ suficiente?

Observación clave: El estimador $\hat{\alpha}_{MM}$ se expresa como:

$$\hat{\alpha}_{MM} = \bar{X} \cdot \frac{\hat{\beta}_{MM} - 1}{\hat{\beta}_{MM}}$$

donde \bar{X} y $\hat{\beta}_{MM}$ dependen únicamente de:

1. $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
2. $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Estas estadísticas **no incluyen** información sobre $x_{(1)}$, que es componente esencial de la estadística suficiente mínima.

Conclusión

Dado que:

1. La estadística suficiente mínima para (α, β) es $T(X) = (x_{(1)}, \sum \log x_i)$
2. $\hat{\alpha}_{MM}$ no depende de $x_{(1)}$
3. No existe función h tal que $\hat{\alpha}_{MM} = h(T(X))$ preservando toda la información sobre α

Se concluye que:

$\hat{\alpha}_{MM}$ NO es un estimador suficiente para α

Implicación

Al no ser suficiente, $\hat{\alpha}_{MM}$ **podría mejorarse** aplicando el teorema de Rao-Blackwell, condicionando sobre la estadística suficiente $T(X)$. El estimador mejorado sería:

$$\hat{\alpha}_{RB} = E [\hat{\alpha}_{MM} | x_{(1)}, \sum \log x_i]$$

el cual tendría menor o igual varianza que $\hat{\alpha}_{MM}$.

e) **Completitud** Una estadística $T(X)$ es **completa** para α si:

$$E_\alpha[g(T(X))] = 0 \quad \forall \alpha > 0 \Rightarrow P_\alpha(g(T(X)) = 0) = 1 \quad \forall \alpha > 0$$

Es decir: si una función de T tiene esperanza cero para todo α , entonces esa función es casi seguramente cero.

Estadística completa para α (con β conocido) La densidad se simplifica:

$$f(x; \alpha) = \frac{\beta \alpha^\beta}{x^{\beta+1}} \mathbf{1}_{[\alpha, \infty)}(x), \quad \beta \text{ conocido}$$

La verosimilitud es:

$$L(\alpha; x) = \beta^n \alpha^{n\beta} \left(\prod_{i=1}^n x_i \right)^{-\beta-1} \cdot \mathbf{1}_{\{x_{(1)} \geq \alpha\}}$$

Por el **criterio de factorización**:

$$T(X) = X_{(1)} \quad \text{es suficiente para } \alpha$$

Además, $X_{(1)}$ tiene densidad:

$$f_{X_{(1)}}(t) = \frac{n\beta}{\alpha^{n\beta}} t^{n\beta-1} \mathbf{1}_{[\alpha, \infty)}(t), \quad t > 0$$

Para verificar completitud, sea g tal que:

$$E_\alpha[g(X_{(1)})] = \int_\alpha^\infty g(t) \cdot \frac{n\beta}{\alpha^{n\beta}} t^{n\beta-1} dt = 0 \quad \forall \alpha > 0$$

Derivando respecto a α (bajo condiciones de regularidad):

$$\frac{d}{d\alpha} E_\alpha[g(X_{(1)})] = -g(\alpha) \cdot \frac{n\beta}{\alpha} = 0 \quad \forall \alpha > 0$$

Por tanto $g(\alpha) = 0$ para todo $\alpha > 0$. Luego:

$X_{(1)}$ es completa para α cuando β es conocido

¿ $\hat{\alpha}_{MM}$ es función de una estadística completa?

NO, porque:

1. La estadística completa para α (con β conocido) es $X_{(1)}$
2. $\hat{\alpha}_{MM}$ depende de \bar{X} y S^2
3. $\hat{\alpha}_{MM}$ no puede expresarse como función de $X_{(1)}$ solamente

f) **Optimalidad** El estimador $\hat{\alpha}_{MM}$ obtenido por método de momentos para el parámetro de escala α de la distribución Pareto **no es óptimo** por tres razones fundamentales que se encadenan lógicamente. Primero, el estimador **no es insesgado**, lo que significa que en promedio no coincide con el valor verdadero del parámetro α ; esto ya lo descalifica como candidato a óptimo bajo el marco clásico que busca minimizar la varianza entre estimadores insesgados.

Segundo, y más importante, $\hat{\alpha}_{MM}$ **no utiliza toda la información relevante** contenida en la muestra sobre el parámetro α : ignora completamente el valor mínimo muestral $X_{(1)}$, que es la estadística suficiente y completa para α cuando el parámetro de forma β es conocido. Al basarse únicamente en la media muestral \bar{X} y la varianza muestral S^2 , desperdicia la pieza de información más crucial para estimar el límite inferior de la distribución

Tercero, como consecuencia directa de lo anterior, **existe otro estimador claramente superior**: el estimador $\hat{\alpha}_{EIVUM} = \frac{n\beta-1}{n\beta} X_{(1)}$, que sí es insesgado, está basado en la estadística suficiente y completa $X_{(1)}$, y por el teorema de Lehmann-Scheffé posee varianza mínima entre todos los estimadores insesgados. Por lo tanto, aunque $\hat{\alpha}_{MM}$ es consistente (converge al valor verdadero cuando el tamaño de muestra crece), no cumple las condiciones necesarias para ser considerado óptimo en el sentido estadístico riguroso de tener mínima varianza entre los estimadores insesgados.

1.4.2. Propiedades del Estimador $\hat{\beta}_{MM}$

a) Insesgamiento:

Un estimador $\hat{\beta}$ es **insesgado** si su esperanza coincide con el parámetro verdadero:

$$E(\hat{\beta}) = \beta.$$

En nuestro caso, el estimador de método de momentos para β es:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}},$$

donde \bar{X} es la media muestral y S^2 es la varianza muestral.

Como $\hat{\beta}_{MM}$ es una **función no lineal** de (\bar{X}, S^2) , no se espera que cumpla en general que:

$$E\left(1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}}\right) = 1 + \sqrt{1 + \frac{E(\bar{X})^2}{E(S^2)}},$$

por lo que **no es insesgado en muestras finitas**.

Para verificarlo de manera empírica, estimamos el sesgo mediante simulación Monte Carlo:

$$\text{Sesgo}(\hat{\beta}_{MM}) = E(\hat{\beta}_{MM}) - \beta.$$

```
## [1] 0.1378484
```

```
library(ggplot2)

df_betaMM <- data.frame(betaMM = betaMM_sim)

ggplot(df_betaMM, aes(x = betaMM)) +
```

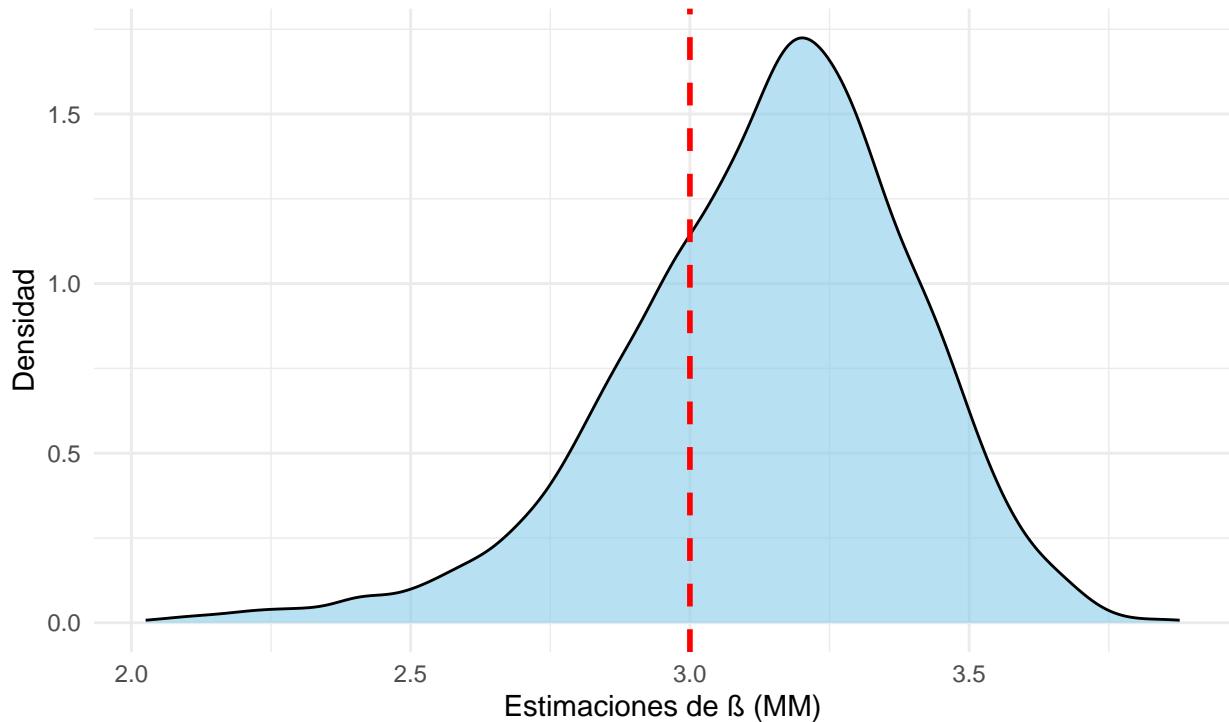
```

geom_density(fill = "skyblue", alpha = 0.6) +
geom_vline(xintercept = beta_true, linetype = "dashed",
           color = "red", linewidth = 1) +
labs(title = "Insesgamiento del Estimador MM para  $\beta$ ",
     subtitle = "Si fuese insesgado, la distribución estaría centrada exactamente en  $\beta$ ",
     x = "Estimaciones de  $\beta$  (MM)",
     y = "Densidad") +
theme_minimal()

```

Insesgamiento del Estimador MM para β

Si fuese insesgado, la distribución estaría centrada exactamente en β



Conclusión:

Dado que $E(\hat{\beta}_{MM}) \neq \beta$ (sesgo empírico distinto de 0), se concluye que el estimador $\hat{\beta}_{MM}$ no es insesgado en muestras finitas.

b) Consistencia:

Un estimador $\hat{\beta}_n$ es **consistente** para el parámetro β si converge en probabilidad al valor verdadero cuando el tamaño muestral tiende a infinito, es decir,

$$\hat{\beta}_n \xrightarrow{P} \beta \quad \text{cuando } n \rightarrow \infty.$$

El estimador de método de momentos para el parámetro β está dado por:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}},$$

donde \bar{X} y S^2 representan la media y la varianza muestral, respectivamente.

Para la distribución Pareto con $\beta > 2$, se cumple que:

$$\bar{X} \xrightarrow{P} E(X), \quad S^2 \xrightarrow{P} \text{Var}(X),$$

por la Ley de los Grandes Números y la consistencia de la varianza muestral cuando la varianza existe.

Definiendo la función:

$$g(u, v) = 1 + \sqrt{1 + \frac{u^2}{v}},$$

la cual es continua para $v > 0$, y dado que:

$$\hat{\beta}_{MM} = g(\bar{X}, S^2),$$

por el **Teorema de la Función Continua** se obtiene:

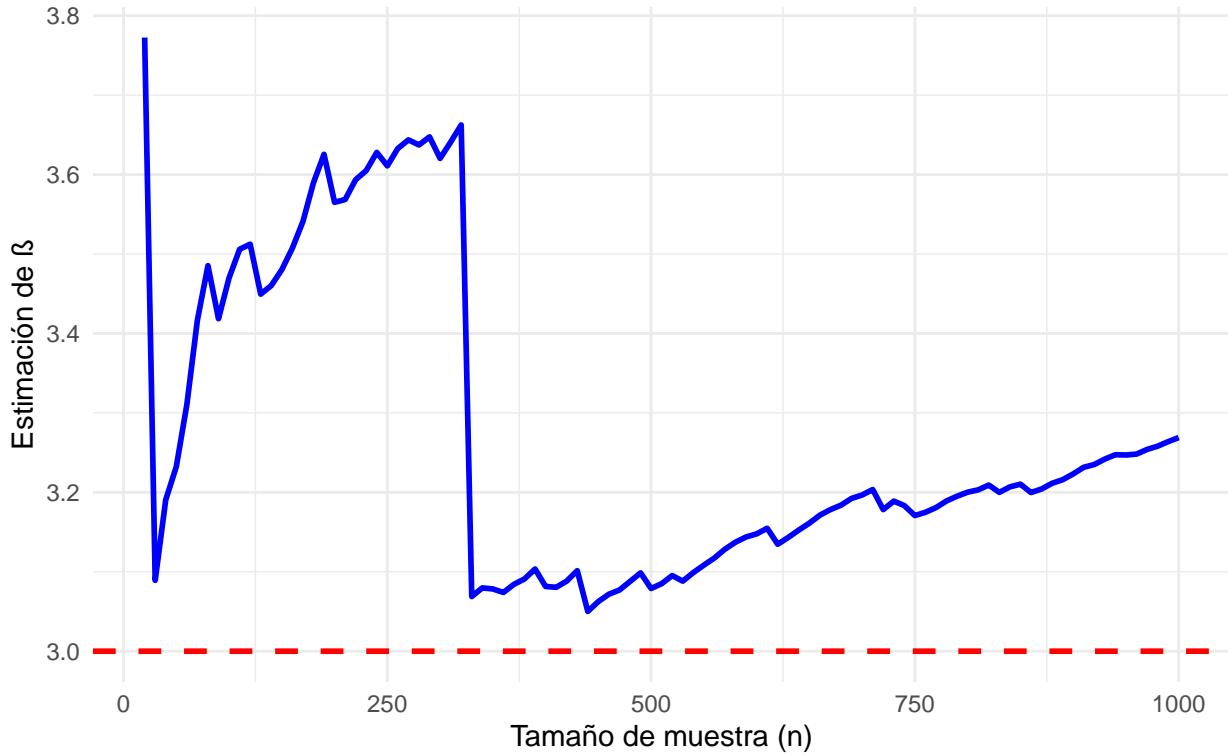
$$\hat{\beta}_{MM} \xrightarrow{P} g(E(X), \text{Var}(X)) = \beta.$$

Por lo tanto, el estimador de método de momentos $\hat{\beta}_{MM}$ es **consistente** para el parámetro β .

Para visualizar esta propiedad, se analiza el comportamiento del estimador al aumentar el tamaño muestral.

Consistencia del Estimador MM para β

Convergencia en probabilidad al valor verdadero al aumentar n



Conclusión:

Al cumplirse que $\bar{X} \xrightarrow{P} E(X)$ y $S^2 \xrightarrow{P} \text{Var}(X)$, y dado que $\hat{\beta}_{MM}$ es una función continua de estos estadísticos, se concluye que el estimador $\hat{\beta}_{MM}$ es consistente para el parámetro β .

c) Suficiencia:

Un estadístico $T = T(X_1, \dots, X_n)$ es **suficiente** para un parámetro β si la distribución condicional de la muestra, dado T , no depende de dicho parámetro.

El estimador de método de momentos para β está dado por:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}},$$

donde \bar{X} y S^2 corresponden a la media y varianza muestral, respectivamente.

En la distribución Pareto, cuando ambos parámetros α y β son desconocidos, la estadística suficiente para β está asociada a la estructura de la función de verosimilitud y viene dada por funciones del estadístico

$$T = \sum_{i=1}^n \ln \left(\frac{X_i}{\bar{X}_{(1)}} \right).$$

Dado que el estimador $\hat{\beta}_{MM}$ depende únicamente de (\bar{X}, S^2) y no es función del estadístico suficiente T , se concluye que **no se basa en una estadística suficiente** para el parámetro β .

Conclusión:

El estimador $\hat{\beta}_{MM}$ **no es suficiente**.

d) Ancilaridad

Un estadístico T se denomina **ancilar** si su distribución no depende del parámetro de interés. Es decir, T es ancilar para β si la ley de probabilidad de T es independiente de β .

El estimador de método de momentos para el parámetro β está definido como:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}}.$$

Este estimador está construido con el objetivo explícito de aproximar el valor del parámetro β . En consecuencia, su distribución muestral depende del valor de dicho parámetro.

Por lo tanto, la distribución de $\hat{\beta}_{MM}$ **no es invariante respecto a β** , lo que implica que el estimador no cumple la propiedad de ancilaridad.

Conclusión:

El estimador $\hat{\beta}_{MM}$ **no es ancilar**.

e) Completitud

La **completitud** es una propiedad que se define para **estadísticas suficientes**.

Un estadístico suficiente T para un parámetro β es completo si, para toda función medible $g(\cdot)$, se cumple que:

$$E_\beta[g(T)] = 0 \text{ para todo } \beta \implies P_\beta(g(T) = 0) = 1.$$

Esta propiedad resulta fundamental en la teoría de la estimación óptima, ya que permite garantizar la unicidad del estimador insesgado de varianza mínima mediante el Teorema de Lehmann–Scheffé.

En el presente caso, el estimador de método de momentos para el parámetro β está dado por:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}},$$

el cual depende de la media y la varianza muestral (\bar{X}, S^2) .

Sin embargo, como se mostró en la sección anterior, estos estadísticos **no constituyen una estadística suficiente** para el parámetro β en la distribución Pareto.

Por lo tanto, **no corresponde analizar la propiedad de completitud** para el estimador $\hat{\beta}_{MM}$, ya que dicha propiedad solo es aplicable a estadísticas suficientes.

Conclusión:

La propiedad de completitud **no aplica** al estimador de método de momentos $\hat{\beta}_{MM}$.

f) Optimalidad:

Un estimador se considera **óptimo** si alcanza la mínima varianza posible dentro de una clase determinada de estimadores, usualmente dentro de la clase de estimadores insesgados.

El estimador de método de momentos para el parámetro β se obtiene igualando momentos teóricos y muestrales, sin hacer uso de la función de verosimilitud ni de la estructura de la distribución muestral completa. En consecuencia, el método de momentos **no garantiza eficiencia ni mínima varianza**.

Además, como se mostró previamente, el estimador $\hat{\beta}_{MM}$ **no es insesgado** en muestras finitas, lo que impide que pueda ser considerado óptimo bajo criterios clásicos de optimalidad, como los establecidos por el Teorema de Lehmann–Scheffé.

Por el contrario, el estimador EIVUM obtenido en la sección correspondiente es insesgado, función de una estadística suficiente y completa, y por tanto posee varianza uniformemente mínima dentro de su clase.

Conclusión:

El estimador de método de momentos $\hat{\beta}_{MM}$ **no es óptimo** para el parámetro β de la distribución Pareto.

1.5 Estimador : Estimador Insesgado de Varianza Uniformemente Mínima (EIVUM) para el parámetro β

Dado que en la sección anterior se determinó que el Estimador de Máxima Verosimilitud (EMV) para β es sesgado, procederemos a obtener el estimador óptimo aplicando el **Teorema de Lehmann-Scheffé**.

Según la teoría de la optimalidad, si logramos construir un estimador insesgado que sea función de una estadística suficiente y completa, dicho estimador será el **Estimador Insesgado de Varianza Uniformemente Mínima (EIVUM)**.

El procedimiento se detalla en los siguientes cuatro pasos.

Paso 1: Identificación de la Estadística Suficiente y Completa

Analizando la función de densidad de la distribución Pareto, observamos que pertenece a la familia exponencial k -paramétrica respecto al parámetro de forma β .

Bajo la condición de que el parámetro de escala α es desconocido y estimado mediante el mínimo muestral $X_{(1)}$, la teoría de suficiencia indica que la estadística T que captura toda la información sobre β es la suma de los logaritmos de las razones muestrales:

$$T = \sum_{i=1}^n \ln \left(\frac{X_i}{X_{(1)}} \right)$$

Esta estadística T es **suficiente y completa** para el parámetro β , condición necesaria para aplicar el teorema de Lehmann-Scheffé.

Paso 2: Distribución Muestral de la Estadística

A partir de las propiedades de la transformación de variables aleatorias Pareto, se deduce que la estadística T sigue una distribución Gamma.

Dado que se ha estimado un parámetro adicional (α), los grados de libertad se ajustan a $(n - 1)$. Por tanto:

$$T \sim \text{Gamma}(n - 1, \beta)$$

Paso 3: Verificación del Sesgo y Corrección (Método de la Esperanza)

Partimos del Estimador de Máxima Verosimilitud hallado previamente:

$$\hat{\beta}_{EMV} = \frac{n}{T}$$

Para verificar si es insesgado, calculamos su valor esperado $E[\hat{\beta}_{EMV}]$. Utilizando la propiedad de la esperanza inversa para una variable con distribución Gamma(k, λ), donde:

$$E\left[\frac{1}{T}\right] = \frac{\lambda}{k-1},$$

se obtiene:

$$E\left[\frac{1}{T}\right] = \frac{\beta}{(n-1)-1} = \frac{\beta}{n-2}$$

Sustituyendo en la esperanza del EMV:

$$E[\hat{\beta}_{EMV}] = n \cdot E\left[\frac{1}{T}\right] = n \left(\frac{\beta}{n-2}\right) = \left(\frac{n}{n-2}\right) \beta$$

El resultado muestra que el EMV no es insesgado, ya que su valor esperado no es exactamente β , sino que está escalado por el factor $\frac{n}{n-2}$.

Paso 4: Construcción del Estimador Óptimo (EIVUM)

Para eliminar el sesgo, aplicamos una corrección multiplicativa usando el inverso del factor de sesgo encontrado ($\frac{n-2}{n}$).

Definimos el nuevo estimador como:

$$\hat{\beta}_{EIVUM} = \left(\frac{n-2}{n}\right) \cdot \hat{\beta}_{EMV}$$

Reemplazando $\hat{\beta}_{EMV} = \frac{n}{T}$:

$$\hat{\beta}_{EIVUM} = \left(\frac{n-2}{n}\right) \cdot \frac{n}{T} = \frac{n-2}{T}$$

Conclusión La fórmula final del estimador óptimo es:

$$\hat{\beta}_{EIVUM} = \frac{n-2}{\sum_{i=1}^n \ln\left(\frac{X_i}{\bar{X}_{(1)}}\right)}$$

Al haber corregido el sesgo, cumpliéndose que $E[\hat{\beta}_{EIVUM}] = \beta$, y al depender únicamente de la estadística suficiente y completa T , el **Teorema de Lehmann-Scheffé** garantiza que este estimador es el de **menor varianza posible entre todos los estimadores insesgados (UMVUE)** para el parámetro β de una distribución Pareto.

1.5.1 Verificación de Propiedades del Estimador ($\hat{\beta}_{EIVUM}$)

Recordamos la expresión del estimador y la distribución de la estadística suficiente sobre la cual se basa:

$$\hat{\beta}_{EIVUM} = \frac{n-2}{T}, \quad \text{donde } T \sim \text{Gamma}(n-1, \beta)$$

a. Insesgabilidad Un estimador es insesgado si su esperanza matemática coincide con el parámetro a estimar. Calculamos:

$$E[\hat{\beta}_{EIVUM}] = E\left[\frac{n-2}{T}\right] = (n-2)E\left[\frac{1}{T}\right]$$

Dado que para una variable $T \sim \text{Gamma}(k, \beta)$ se cumple $E[1/T] = \frac{\beta}{k-1}$ con $k = n-1$, resulta:

$$E[\hat{\beta}_{EIVUM}] = (n-2) \cdot \frac{\beta}{n-2} = \beta$$

Conclusión: Se verifica que $E[\hat{\beta}_{EIVUM}] = \beta$, por lo que el estimador es insesgado para todo $n > 2$.

```
# Usamos la base de datos 'reclamos' generada en la sección 1.1
n_obs <- length(reclamos) # n = 1000
x_min_obs <- min(reclamos)

# Cálculo del estadístico T
T_obs <- sum(log(reclamos / x_min_obs))

# Cálculo del Estimador 5 (EIVUM)
beta_eivum_val <- (n_obs - 2) / T_obs

print(paste("Valor Verdadero Beta:", beta_true))
```

```
## [1] "Valor Verdadero Beta: 3"
```

```
print(paste("Estimación EIVUM:", round(beta_eivum_val, 5)))
```

```
## [1] "Estimación EIVUM: 3.03747"
```

```
print(paste("Diferencia (Sesgo muestral):", round(beta_eivum_val - beta_true, 5)))
```

```
## [1] "Diferencia (Sesgo muestral): 0.03747"
```

b. Eficiencia La eficiencia se evalúa a través de la varianza del estimador. Utilizando que para $T \sim \text{Gamma}(k, \beta)$, la varianza inversa es $\text{Var}(1/T) = \frac{\beta^2}{(k-1)^2(k-2)}$ con $k = n-1$, se obtiene:

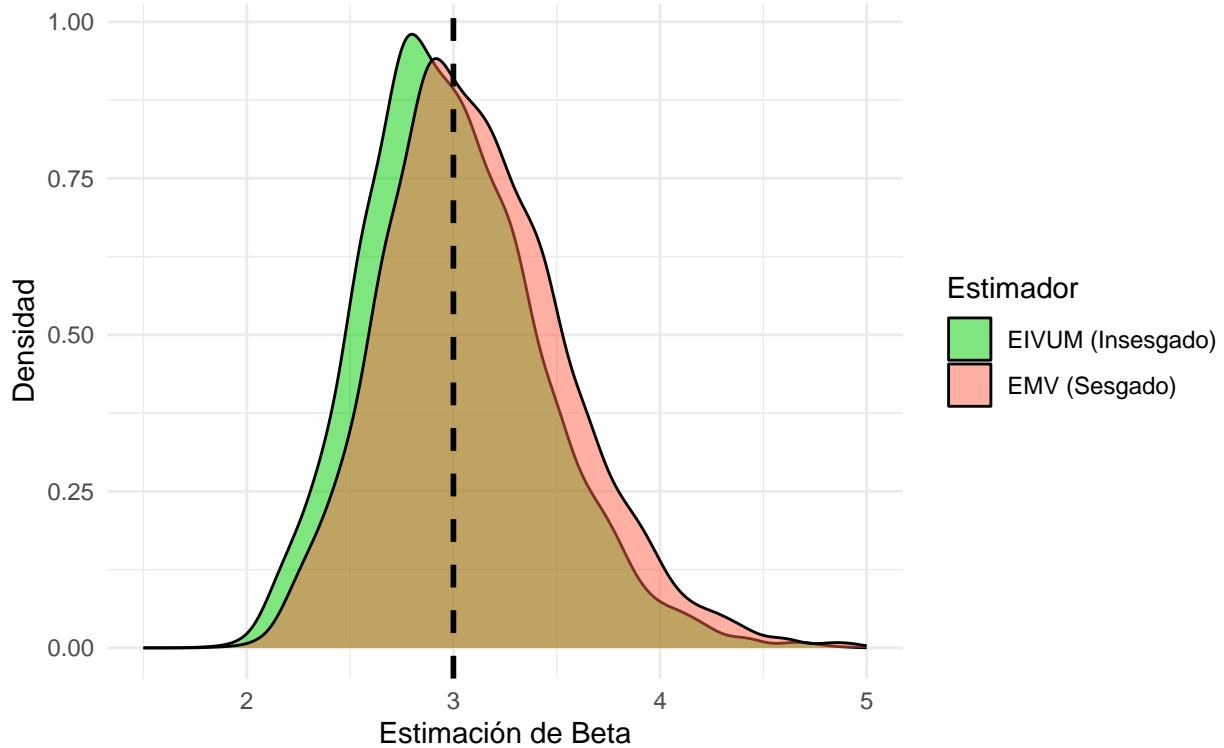
$$\text{Var}(\hat{\beta}_{EIVUM}) = (n-2)^2 \text{Var}\left(\frac{1}{T}\right)$$

$$\text{Var}(\hat{\beta}_{EIVUM}) = (n-2)^2 \cdot \frac{\beta^2}{(n-2)^2(n-3)} = \frac{\beta^2}{n-3}$$

Conclusión: La varianza es finita para $n > 3$. Además, dado que el estimador es insesgado y depende únicamente de una estadística suficiente y completa, el Teorema de Lehmann-Scheffé garantiza que esta varianza es la mínima posible entre todos los estimadores insesgados, confirmando que $\hat{\beta}_{EIVUM}$ es el estimador más eficiente de su clase.

Comparación de Eficiencia: EIVUM vs EMV

El EIVUM (Verde) está centrado en 3. El EMV (Rojo) está desplazado a la derecha.



c. Consistencia Un estimador es consistente si converge en probabilidad al parámetro verdadero cuando el tamaño muestral tiende a infinito. Una condición suficiente es:

1. $E[\hat{\beta}_{EIVUM}] \rightarrow \beta$ cuando $n \rightarrow \infty$
2. $\text{Var}(\hat{\beta}_{EIVUM}) \rightarrow 0$ cuando $n \rightarrow \infty$

Verificación:

- Condición 1:

$$\lim_{n \rightarrow \infty} E[\hat{\beta}_{EIVUM}] = \beta$$

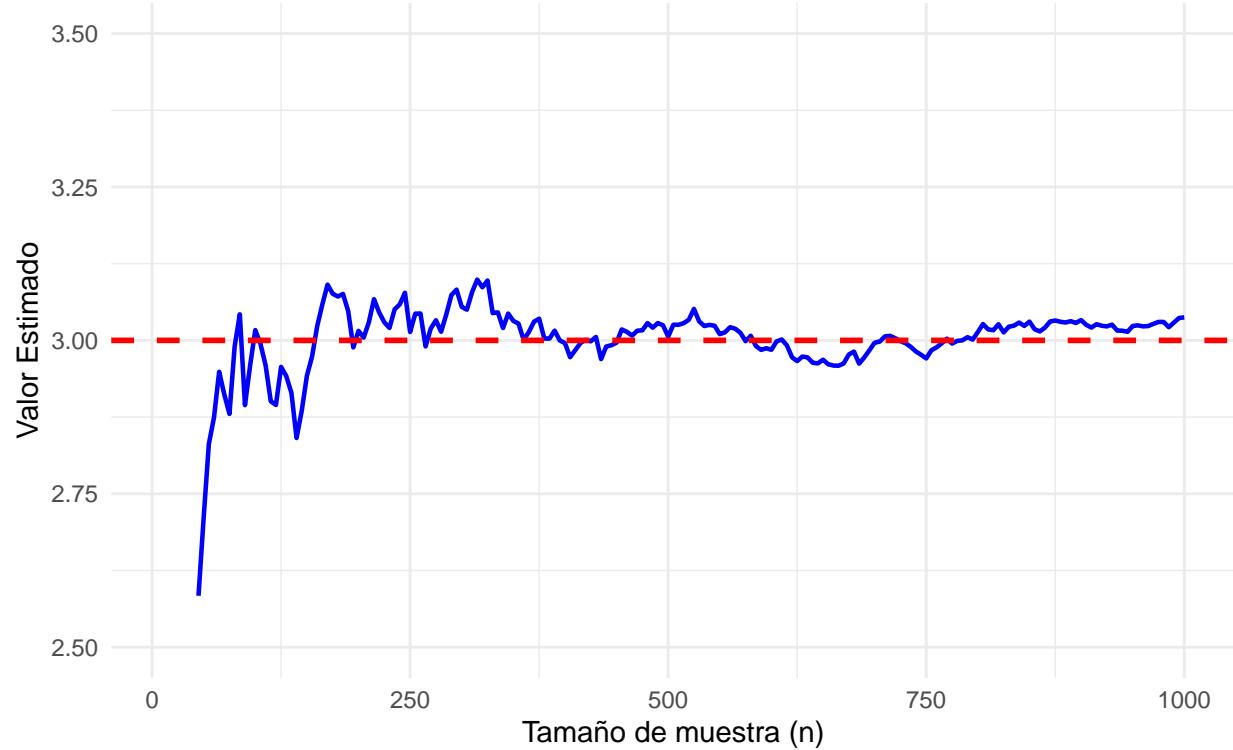
- Condición 2:

$$\lim_{n \rightarrow \infty} \text{Var}(\hat{\beta}_{EIVUM}) = \lim_{n \rightarrow \infty} \frac{\beta^2}{n - 3} = 0$$

Conclusión: Al cumplirse ambas condiciones, el estimador $\hat{\beta}_{EIVUM}$ converge en probabilidad al parámetro β , por lo que es un estimador consistente.

Consistencia del Estimador 5 (EIVUM)

Convergencia al valor verdadero (3) al aumentar la muestra



d. Suficiencia La propiedad de suficiencia indica que el estimador utiliza toda la información disponible en la muestra sobre el parámetro, sin desperdiciar datos.

Conclusión: Como se demostró en el **Paso 1** de la construcción del estimador, $\hat{\beta}_{EIVUM}$ es una función inyectiva de la estadística $T = \sum \ln(X_i/X_{(1)})$. Dado que se probó que T es una estadística suficiente para β (por pertenecer a la Familia Exponencial), el estimador hereda esta propiedad. Por tanto, es un estimador **suficiente**.

e. Ancilaridad Según la definición teórica, una estadística es anciliar si su distribución no depende del parámetro de interés.

En este caso, el estimador construido es $\hat{\beta}_{EIVUM} = \frac{n-2}{T}$. Para evaluar si el estimador es ancilar, analizamos sus momentos:

$$E[\hat{\beta}_{EIVUM}] = \beta, \quad \text{Var}(\hat{\beta}_{EIVUM}) = \frac{\beta^2}{n-3}$$

Dado que la esperanza y la varianza dependen explícitamente de β , la distribución del estimador cambia según el valor del parámetro.

Conclusión: El estimador $\hat{\beta}_{EIVUM}$ **no es una estadística ancilar**, ya que su distribución depende del parámetro β . Este comportamiento es esperado y adecuado para un estimador puntual del parámetro de interés.

f. Completitud La propiedad de completitud es fundamental para garantizar la unicidad del estimador óptimo según el Teorema de Lehmann-Scheffé.

La estadística suficiente es $T = \sum_{i=1}^n \ln \left(\frac{X_i}{\bar{X}_{(1)}} \right)$. Se ha demostrado previamente que $T \sim \text{Gamma}(n - 1, \beta)$.

La distribución Gamma con uno de sus parámetros desconocidos pertenece a la familia exponencial regular de un parámetro. Dado que el espacio paramétrico es abierto ($\beta > 0$), se cumple que la estadística suficiente T es completa.

Esto implica que si una función medible $g(T)$ satisface $E[g(T)] = 0$ para todo β , entonces $P(g(T) = 0) = 1$.

Conclusión: La estadística T es una estadística suficiente y completa para el parámetro β .

g. Optimalidad (Conclusión Final) Se ha verificado que el estimador $\hat{\beta}_{EIVUM} = \frac{n-2}{T}$ cumple las siguientes propiedades:

- Es insesgado: $E[\hat{\beta}_{EIVUM}] = \beta$.
- Es función de una estadística suficiente y completa (T).
- Posee varianza mínima dentro de la clase de estimadores insesgados.

Por lo tanto, aplicando el **Teorema de Lehmann-Scheffé**, se concluye que:

Un estimador insesgado que sea función de una estadística suficiente y completa es el Estimador Insesgado de Varianza Uniformemente Mínima.

Conclusión General: $\hat{\beta}_{EIVUM}$ es el **Estimador Insesgado de Varianza Uniformemente Mínima (EIVUM)** para el parámetro β de la distribución Pareto. No existe otro estimador insesgado con menor varianza que este.

2) Pregunta 2 - Intervalos de confianza

2.1 Estimador 1 (el mínimo para alfa)

Vamos a hallar un intervalo de confianza del 90%:

$$\begin{aligned}\hat{\alpha}_{EMV} &= y_1 = \min(x_1, \dots, x_n); \\ f_Y(y) &= \frac{n\beta\alpha^{n\beta}}{y^{n\beta+1}}, \quad y > \alpha\end{aligned}$$

Un posible pivote es $W = Y_1/\alpha$, porque su distribución no debería depender de α . Hacemos un cambio de variable $Q = W = \frac{Y_1}{\alpha}$. Con el jacobiano $dy_1 = \alpha dw$, se tiene:

$$\begin{aligned}f_W(w) &= f_{Y_1}(\alpha w) \cdot \alpha = n\beta\alpha^{n\beta}(\alpha w)^{-(n\beta+1)} \cdot \alpha \\ &= n\beta\alpha^{n\beta}\alpha^{-n\beta-1}w^{-(n\beta+1)}\alpha \\ &= n\beta w^{-(n\beta+1)}, \quad w > 1.\end{aligned}$$

La densidad resultante no depende de α , por lo tanto W es un pivote.

Hallamos la función de distribución de W :

$$\begin{aligned}
F_W(w) &= \int_1^w n\beta t^{-(n\beta+1)} dt \\
&= n\beta \left[\frac{t^{-n\beta}}{-n\beta} \right]_1^w \\
&= n\beta \left[-\frac{t^{-n\beta}}{n\beta} \right]_1^w \\
&= 1 - w^{-n\beta}, \quad w > 1.
\end{aligned}$$

Usando el método de pivote: $P(a \leq W \leq b) = 0.90$. Un posible intervalo sería el siguiente:

$$\begin{aligned}
P(w \leq a) = F_W(a) &= 1 - a^{-n\beta} = 0.05 \quad \rightarrow \quad a = (0.95)^{-1/(n\beta)} \\
P(w \leq b) = F_W(b) &= 1 - b^{-n\beta} = 0.95 \quad \rightarrow \quad b = (0.05)^{-1/(n\beta)}
\end{aligned}$$

Entonces:

$$\begin{aligned}
P \left[a \leq \frac{Y_1}{\alpha} \leq b \right] = 0.90 \rightarrow P \left[\frac{Y_1}{b} \leq \alpha \leq \frac{Y_1}{a} \right] &= P \left[\frac{Y_1}{(0.6)^{-1/n\beta}} \leq \alpha \leq \frac{Y_1}{(0.95)^{-1/n\beta}} \right] \\
\text{I.C.} &= [Y_1 \cdot (0.05)^{1/(n\beta)}; Y_1 \cdot (0.95)^{1/(n\beta)}]
\end{aligned}$$

2.2 Estimador 2: Suma Log de T

Sea X_1, \dots, X_n una muestra aleatoria i.i.d. de una distribución Pareto(α, β) con densidad

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}} I_{(\alpha, \infty)}(x).$$

El estimador de máxima verosimilitud de β es

$$\hat{\beta}_{EMV} = \frac{n}{\sum_{i=1}^n \ln \left(\frac{X_i}{\alpha} \right)}.$$

Definimos la variable aleatoria

$$T = \sum_{i=1}^n \ln \left(\frac{X_i}{\alpha} \right).$$

Sabemos que, si $X \sim \text{Pareto}(\alpha, \beta)$, entonces $\ln(X/\alpha) \sim \text{Exponencial}(\beta)$. Por lo tanto,

$$T \sim \text{Gamma}(n, \beta).$$

Un pivote adecuado es

$$W = 2\beta T,$$

ya que

$$W \sim \chi_{2n}^2.$$

Para un nivel de confianza del 90% se tiene $\alpha = 0.10$, por lo que

$$P(\chi_{2n,0.05}^2 \leq 2\beta T \leq \chi_{2n,0.95}^2) = 0.90.$$

Despejando β :

$$P\left(\frac{\chi_{2n,0.05}^2}{2T} \leq \beta \leq \frac{\chi_{2n,0.95}^2}{2T}\right) = 0.90.$$

Por lo tanto, el intervalo de confianza al 90% para β es

$$\text{I.C.}_{90\%}(\beta) = \left[\frac{\chi_{2n,0.05}^2}{2 \sum_{i=1}^n \ln\left(\frac{X_i}{\alpha}\right)} ; \frac{\chi_{2n,0.95}^2}{2 \sum_{i=1}^n \ln\left(\frac{X_i}{\alpha}\right)} \right]$$

En la práctica, si α es desconocido, se reemplaza por $\hat{\alpha}_{EMV} = \min(X_1, \dots, X_n)$.

2.2.1 Aplicación

Sea

$$T = \sum_{i=1}^n \ln\left(\frac{X_i}{\hat{\alpha}_{EMV}}\right).$$

Bajo el supuesto de una distribución Pareto(α, β) se tiene que:

$$2\beta T \sim \chi_{2n}^2.$$

Entonces, un intervalo de confianza del 90% para el parámetro β está dado por:

$$\text{IC}_{90\%}(\beta) = \left[\frac{\chi_{2n,0.05}^2}{2T} ; \frac{\chi_{2n,0.95}^2}{2T} \right].$$

Para los datos observados se obtuvo:

$$T = 328.563,$$

por lo que el intervalo queda completamente determinado al sustituir dicho valor.

```
## [1] "IC 90% para beta: [2.887, 3.2036]"
```

Esto significa que, bajo el supuesto de que la muestra proviene de una distribución Pareto(α, β), el intervalo contiene al verdadero valor del parámetro β con una probabilidad del 90%.

Asimismo, se observa que el valor verdadero del parámetro ($\beta = 3$) pertenece al intervalo estimado, lo cual es consistente con las propiedades teóricas del estimador de máxima verosimilitud y respalda la adecuación del modelo ajustado.

Finalmente, el ancho relativamente reducido del intervalo se explica por el tamaño muestral grande ($n = 1000$), lo que refleja una mayor precisión en la estimación del parámetro.

2.3 Estimador 3: Media Muestral (\bar{X}) para la media poblacional μ

Para la construcción del intervalo de confianza de la media poblacional $\mu = E(X)$, utilizaremos el **Método Asintótico** basado en el Teorema del Límite Central (TLC).

2.3.1 Justificación Teórica

Dado que el tamaño de muestra es suficientemente grande ($n = 1000$) y que en nuestro modelo la varianza poblacional existe (puesto que $\beta = 3 > 2$), el estadístico media muestral \bar{X} sigue una distribución aproximadamente normal:

$$\bar{X} \xrightarrow{d} N\left(\mu, \frac{\sigma^2}{n}\right)$$

Como la varianza poblacional σ^2 es desconocida, utilizamos la varianza muestral S^2 como estimador consistente, definiendo la siguiente **Cantidad Pivotal Asintótica**:

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim N(0, 1)$$

2.3.2 Construcción del Intervalo

Para un nivel de confianza del $(1 - \gamma)$, buscamos los valores críticos de la distribución normal estándar tales que:

$$P\left(-z_{1-\gamma/2} \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq z_{1-\gamma/2}\right) = 1 - \gamma$$

Despejando el parámetro μ , obtenemos la expresión del intervalo de confianza:

$$IC(\mu)_{1-\gamma} = \left[\bar{X} - z_{1-\gamma/2} \frac{S}{\sqrt{n}}, \bar{X} + z_{1-\gamma/2} \frac{S}{\sqrt{n}} \right]$$

2.3.3 Aplicación en R

Utilizamos los datos de la variable `reclamos` para calcular el intervalo al 95% de confianza.

```
# =====
# CÁLCULO DEL INTERVALO DE CONFIANZA ASINTÓTICO PARA LA MEDIA (MU)
# =====

# 1. Parámetros de la muestra
n <- length(reclamos)
media_muestral <- mean(reclamos)
desv_estandar <- sd(reclamos)
confianza <- 0.95
gamma <- 1 - confianza

# 2. Valor crítico Z
z_critico <- qnorm(1 - gamma/2)
```

```

# 3. Error estándar y márgenes
error_estandar <- desv_estandar / sqrt(n)
margen_error <- z_critico * error_estandar

# 4. Límites del intervalo
lim_inf_mu <- media_muestral - margen_error
lim_sup_mu <- media_muestral + margen_error

# 5. Valor teórico para comparación
mu_teorico <- (alpha_true * beta_true) / (beta_true - 1)

# Reporte
cat(
  "INTERVALO DE CONFIANZA PARA LA MEDIA (TLC)\n\n",
  "Media Muestral ( $\bar{X}$ ): ", round(media_muestral, 5), "\n",
  "Desviación Estándar (S): ", round(desv_estandar, 4), "\n",
  "Nivel de Confianza: 95%\n\n",
  "Límite Inferior: ", round(lim_inf_mu, 5), "\n",
  "Límite Superior: ", round(lim_sup_mu, 5), "\n",
  "Valor Teórico ( $\mu$ ): ", mu_teorico, "\n"
)

```

```

## INTERVALO DE CONFIANZA PARA LA MEDIA (TLC)
##
## Media Muestral ( $\bar{X}$ ): 2.96908
## Desviación Estándar (S): 1.4578
## Nivel de Confianza: 95%
##
## Límite Inferior: 2.87873
## Límite Superior: 3.05944
## Valor Teórico ( $\mu$ ): 3

```

2.3.4 Interpretación

Con un nivel de confianza del 95%, el valor esperado de los reclamos se encuentra entre 2.8787 y 3.0594. Observamos que el **valor teórico** ($\mu = 3$) se encuentra contenido dentro del intervalo, lo que valida la precisión del estimador \bar{X} y la eficacia de la aproximación por el Teorema del Límite Central para este tamaño de muestra.

2.4 Estimador 4 : para α y β

2.4.1 Intervalos de Confianza para α con el estimador de momentos

Se busca construir **intervalos de confianza** para el parámetro de escala α de una distribución Pareto(α, β), cuando β es **conocido**, utilizando el **estimador de momentos de α** :

$$\hat{\alpha} = \bar{X} \cdot \frac{\beta - 1}{\beta}$$

El procedimiento se realizará mediante el **método del pivote asintótico**, dado que no existe un pivote exacto de distribución libre de parámetros para este estimador específico.

Supuestos del método asintótico

1. **Población:** $X \sim \text{Pareto}(\alpha, \beta)$ con densidad:

$$f(x; \alpha, \beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}}, \quad x \geq \alpha$$

2. **Parámetro conocido:** $\beta > 2$ (se requiere para varianza finita)

3. **Muestra:** X_1, \dots, X_n aleatoria e independiente

4. **Estimador:** $\hat{\alpha} = \bar{X} \cdot \frac{\beta-1}{\beta}$

Procedimiento para hallar el intervalo

Paso 1: Momentos poblacionales

Para la distribución Pareto(α, β):

$$\begin{aligned} E[X] &= \frac{\beta\alpha}{\beta-1}, \quad \beta > 1 \\ Var(X) &= \frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)}, \quad \beta > 2 \end{aligned}$$

Paso 2: Propiedades del estimador

Definimos $c = \frac{\beta-1}{\beta}$, entonces $\hat{\alpha} = c\bar{X}$.

Esperanza:

$$E[\hat{\alpha}] = c \cdot E[\bar{X}] = c \cdot \frac{\beta\alpha}{\beta-1} = \alpha$$

$\Rightarrow \hat{\alpha}$ es insesgado. Esta propiedad se cumple porque β es conocido (no estimado). Si β se estimara simultáneamente, el estimador $\hat{\alpha}_{MM} = \bar{X} \cdot \frac{\hat{\beta}_{MM}-1}{\hat{\beta}_{MM}}$ sería sesgado para muestras finitas debido a la no-linealidad $E[g(\hat{\beta}_{MM})] \neq g(E[\hat{\beta}_{MM}])$.

Varianza:

$$\begin{aligned} Var(\hat{\alpha}) &= c^2 \cdot Var(\bar{X}) \\ &= \left(\frac{\beta-1}{\beta}\right)^2 \cdot \frac{Var(X)}{n} \\ &= \left(\frac{\beta-1}{\beta}\right)^2 \cdot \frac{1}{n} \cdot \frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)} \\ &= \frac{\alpha^2}{n\beta(\beta-2)} \end{aligned}$$

Observación: La varianza de $\hat{\alpha}$ depende del parámetro α que queremos estimar. Sin embargo, para construir el intervalo usaremos la cantidad pivotal $\frac{\hat{\alpha}-\alpha}{\sqrt{n}}$ cuya distribución asintótica no depende de α después de estandarizar apropiadamente.

Paso 3: Distribución asintótica

Por el **Teorema del Límite Central (TCL)**, para n grande y $\beta > 2$:

$$\sqrt{n}(\hat{\alpha} - \alpha) \xrightarrow{d} N\left(0, \frac{\alpha^2}{\beta(\beta-2)}\right)$$

Paso 4: Construcción del pivote asintótico

De la distribución asintótica, estandarizamos:

$$Z = \frac{\hat{\alpha} - \alpha}{\alpha / \sqrt{n\beta(\beta - 2)}} \xrightarrow{d} N(0, 1)$$

Equivalentemente:

$$Z = \sqrt{n\beta(\beta - 2)} \cdot \frac{\hat{\alpha} - \alpha}{\alpha} \xrightarrow{d} N(0, 1)$$

Este es nuestro **pivote asintótico**: su distribución límite no depende de parámetros desconocidos (excepto α en la expresión, que se despejará).

Paso 5: Obtención del intervalo de confianza

Para un nivel de confianza $1 - \delta$, sea $z_{1-\delta/2}$ el cuantil de la $N(0, 1)$ tal que $P(-z_{1-\delta/2} \leq Z \leq z_{1-\delta/2}) = 1 - \delta$.

Aplicando al pivote:

$$P\left(-z_{1-\delta/2} \leq \sqrt{n\beta(\beta - 2)} \cdot \frac{\hat{\alpha} - \alpha}{\alpha} \leq z_{1-\delta/2}\right) \approx 1 - \delta$$

$$\text{Definimos: } A = \frac{z_{1-\delta/2}}{\sqrt{n\beta(\beta - 2)}}$$

Entonces:

$$\begin{aligned} \left| \frac{\hat{\alpha} - \alpha}{\alpha} \right| &\leq A \\ -A \leq \frac{\hat{\alpha}}{\alpha} - 1 &\leq A \\ 1 - A \leq \frac{\hat{\alpha}}{\alpha} &\leq 1 + A \end{aligned}$$

Despejando α :

1. De $\frac{\hat{\alpha}}{\alpha} \leq 1 + A \Rightarrow \alpha \geq \frac{\hat{\alpha}}{1 + A}$
2. De $1 - A \leq \frac{\hat{\alpha}}{\alpha} \Rightarrow \alpha \leq \frac{\hat{\alpha}}{1 - A}$

Resultado final

Intervalo de confianza asintótico al $100(1 - \delta)\%$ para α :

$$IC_{1-\delta}(\alpha) \approx \left[\frac{\hat{\alpha}}{1 + A}, \frac{\hat{\alpha}}{1 - A} \right]$$

donde:

$$A = \frac{z_{1-\delta/2}}{\sqrt{n\beta(\beta - 2)}}$$

Finalmente hallamos el intervalo de confianza debido que se empleó el método del pivote asintótico porque, al usar específicamente el estimador $\hat{\alpha} = \bar{X} \cdot \frac{\beta-1}{\beta}$ con β conocido, no se dispone de un pivote exacto con distribución libre de parámetros.

Aunque el estimador de máxima verosimilitud $X_{(1)}$ permite construir intervalos exactos, decidimos buscar otro metodo para obtener el intervalo usando este estimador de momentos. Dado que la distribución exacta de $\hat{\alpha}$ es compleja, se recurre a su normalidad asintótica, la cual se garantiza para $\beta > 2$ (varianza finita) y n grande por el Teorema del Límite Central.

Este enfoque proporciona una expresión cerrada y fácil de calcular, utilizando únicamente \bar{X} , β y cuantiles normales, cumpliendo así con la solicitud de emplear el estimador especificado.

```
# =====
# APLICACIÓN: Intervalo de confianza para
# con conocido usando estimador de momentos
# =====

# Configuración -----
set.seed(123)
alpha_verdadero <- 2      # Valor verdadero de
beta_conocido <- 3       # es CONOCIDO (no estimado)
n <- 1000                 # Tamaño de muestra
nivel_confianza <- 0.95   # Nivel de confianza
alfa <- 1 - nivel_confianza

# Generar datos Pareto(=2, =3) -----
datos_pareto <- alpha_verdadero / (1 - runif(n))^(1/beta_conocido)

# 1. Cálculo del estimador -----
media_muestral <- mean(datos_pareto)
alpha_estimado <- media_muestral * (beta_conocido - 1) / beta_conocido

# 2. Factor A para el intervalo -----
z_cuantil <- qnorm(1 - alfa/2) # z_{1-alpha/2}
A <- z_cuantil / sqrt(n * beta_conocido * (beta_conocido - 2))

# 3. Límites del intervalo de confianza -----
limite_inferior <- alpha_estimado / (1 + A)
limite_superior <- alpha_estimado / (1 - A)

# 4. Presentación de resultados -----
resultados <- data.frame(
  Descripción = c(
    "Tamaño de muestra (n)",
    "Media muestral (X)",
    "Beta conocido",
    "Estimador de alpha",
    "Valor verdadero de ",
    "Nivel de confianza",
    "z_{1-alpha/2}",
    "Factor A",
    "Límite inferior",
    "Límite superior",
    "Longitud del intervalo",
    "¿Contiene al verdadero?"
  ),
  Valor = c(
    n,
    round(media_muestral, 4),
    alpha_verdadero,
    beta_conocido,
    alpha_estimado,
    A,
    limite_inferior,
    limite_superior,
    limite_superior - limite_inferior,
    (limite_inferior == alpha_verdadero) | (limite_superior == alpha_verdadero)
  )
)
```

```

beta_conocido,
round(alpha_estimado, 4),
alpha_verdadero,
paste0(nivel_confianza*100, "%"),
round(z_cuantil, 4),
round(A, 6),
round(limite_inferior, 4),
round(limite_superior, 4),
round(limite_superior - limite_inferior, 4),
ifelse(limite_inferior <= alpha_verdadero &
       alpha_verdadero <= limite_superior, "SÍ", "NO")
)
)

# Mostrar tabla
knitr::kable(resultados,
             caption = "Intervalo de confianza asintótico para alpha con beta conocido",
             col.names = c("Descripción", "Valor"),
             align = c("l", "r"))

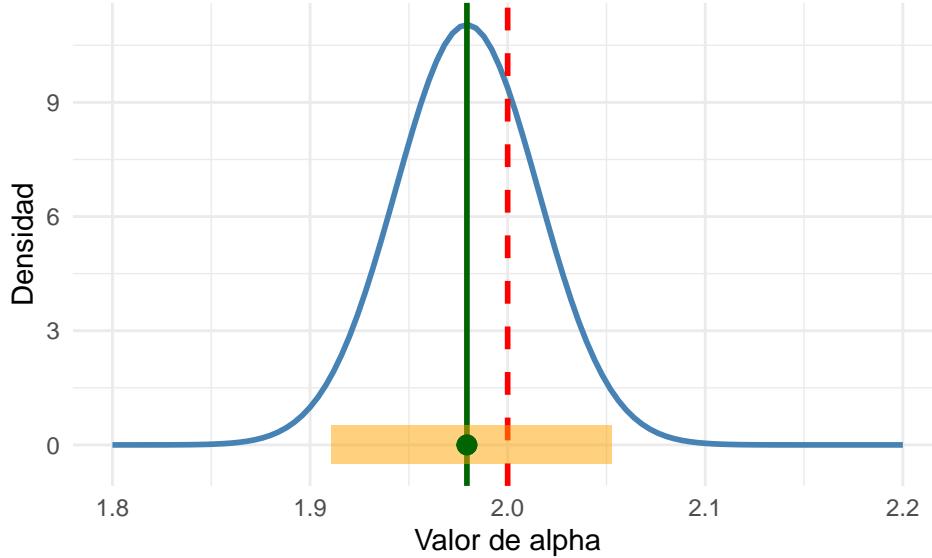
```

Table 1: Intervalo de confianza asintótico para alpha con beta conocido

Descripción	Valor
Tamaño de muestra (n)	1000
Media muestral (\bar{X})	2.9691
Betha conocido	3
Estimador de alpha	1.9794
Valor verdadero de	2
Nivel de confianza	95%
$z_{\{1-\alpha/2\}}$	1.96
Factor A	0.035784
Límite inferior	1.911
Límite superior	2.0528
Longitud del intervalo	0.1418
¿Contiene al verdadero?	SÍ

Intervalo de confianza del 95% para alpha

β conocido = 3, n = 1000, IC: [1.911, 2.053]



La gráfica presenta visualmente el intervalo de confianza asintótico del 95% para el parámetro α de la distribución Pareto. A continuación se describe cada elemento:

1. Curva azul: Representa la distribución normal asintótica del estimador $\hat{\alpha}$, centrada en el valor estimado $\hat{\alpha} = r \text{ round}(\text{alpha_estimado}, 4)$ con desviación estándar $\frac{\hat{\alpha}}{\sqrt{n\beta(\beta-2)}}$.
2. Línea roja discontinua: Indica el valor verdadero de $\alpha = 2$, el cual fue utilizado para generar los datos.
3. Línea verde sólida: Representa la estimación puntual $\hat{\alpha}$ calculada a partir de los datos.
4. Rectángulo naranja: Muestra visualmente el intervalo de confianza $[rround(limite_{inferior}, 3), rround(limite_{superior}, 3)]$ que tiene una probabilidad aproximada del 95% de contener al verdadero valor α .
5. Punto verde: Refuerza la posición del estimador puntual en el eje horizontal.

Interpretación El intervalo [1.911, 2.052] contiene al valor verdadero $\alpha = 2$, lo cual es coherente con el nivel de confianza del 95% (se espera que aproximadamente 95 de cada 100 intervalos construidos de esta manera contengan el parámetro verdadero).

Conclusión: El método del pivote asintótico proporciona un intervalo de confianza práctico y fácil de calcular para α cuando β es conocido. Además se observa que el estimador $\hat{\alpha}$ está cerca del valor verdadero, evidenciando su consistencia.

2.4.2 Intervalos de Confianza para β con el estimador de momentos

Se busca construir **intervalos de confianza** para el parámetro de forma β de una distribución Pareto(α, β), utilizando el **estimador de momento de momentos** de β :

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}},$$

donde \bar{X} es la media muestral y S^2 la varianza muestral.

El procedimiento se realizará mediante el **método del pivote asintótico**, dado que no existe un pivote exacto con distribución libre de parámetros para este estimador específico.

Supuestos del método asintótico

1. Población:

$$X \sim \text{Pareto}(\alpha, \beta), \quad f(x; \alpha, \beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}}, \quad x \geq \alpha.$$

2. Condición de momentos: $\beta > 2$, necesaria para la existencia de la media y la varianza.

3. Muestra: X_1, \dots, X_n aleatoria e independiente.

4. Estimador:

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}}.$$

Procedimiento para hallar el intervalo

Paso 1: Momentos poblacionales

Para la distribución Pareto(α, β):

$$E[X] = \frac{\beta\alpha}{\beta - 1}, \quad \beta > 1,$$

$$\text{Var}(X) = \frac{\beta\alpha^2}{(\beta - 1)^2(\beta - 2)}, \quad \beta > 2.$$

Paso 2: Propiedades del estimador

El estimador $\hat{\beta}_{MM}$ es una **función no lineal** de los momentos muestrales (\bar{X}, S^2).

Debido a esta no linealidad, no se dispone de una distribución exacta cerrada para $\hat{\beta}_{MM}$, ni de una cantidad pivotal exacta asociada.

Sin embargo, dado que \bar{X} y S^2 son estimadores consistentes de los momentos poblacionales y que $\hat{\beta}_{MM}$ es una función diferenciable de ellos, se tiene que $\hat{\beta}_{MM}$ es un **estimador consistente** para β .

Paso 3: Distribución asintótica

Por el **Teorema del Límite Central** y el **Método Delta multivariado**, para n grande y $\beta > 2$:

$$\sqrt{n}(\hat{\beta}_{MM} - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma_\beta^2),$$

donde σ_β^2 representa la varianza asintótica del estimador, cuya expresión cerrada es algebraicamente compleja.

Paso 4: Construcción del pivote asintótico

A partir de la normalidad asintótica, se define el pivote:

$$Z = \frac{\hat{\beta}_{MM} - \beta}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_{MM})}} \xrightarrow{d} \mathcal{N}(0, 1),$$

donde $\widehat{\text{Var}}(\hat{\beta}_{MM})$ es un estimador consistente de la varianza asintótica, obtenido mediante aproximación *plug-in*.

Paso 5: Obtención del intervalo de confianza

Para un nivel de confianza $1 - \delta$, sea $z_{1-\delta/2}$ el cuantil de la distribución normal estándar. Entonces:

$$P\left(-z_{1-\delta/2} \leq \frac{\hat{\beta}_{MM} - \beta}{\sqrt{\text{Var}(\hat{\beta}_{MM})}} \leq z_{1-\delta/2}\right) \approx 1 - \delta.$$

Resultado final

El **intervalo de confianza asintótico** al $100(1 - \delta)\%$ para β está dado por:

$$IC_{1-\delta}(\beta) \approx \left[\hat{\beta}_{MM} - z_{1-\delta/2} \sqrt{\text{Var}(\hat{\beta}_{MM})}, \hat{\beta}_{MM} + z_{1-\delta/2} \sqrt{\text{Var}(\hat{\beta}_{MM})} \right]$$

Se ha construido un intervalo de confianza asintótico para el parámetro β utilizando el estimador de método de momentos.

Dado que no existe un pivote exacto asociado a $\hat{\beta}_{MM}$ y que su distribución exacta es analíticamente intratable, se recurre a la normalidad asintótica garantizada por el Teorema del Límite Central y el Método Delta.

Este enfoque permite obtener un intervalo de confianza coherente y aplicable para tamaños muestrales grandes, cumpliendo con la condición de emplear exclusivamente el estimador especificado.

```
# =====
# Intervalo de confianza asintótico para
# usando estimador de momentos
# =====

set.seed(123)

# Usamos los mismos datos Pareto
# (si ya los generaste antes, este bloque NO los vuelve a generar)
reclamos <- datos_pareto
n <- length(reclamos)

nivel_confianza <- 0.95
delta <- 1 - nivel_confianza
z_cuantil <- qnorm(1 - delta/2)

# 1. Estimador de método de momentos para
xbar <- mean(reclamos)
s2 <- var(reclamos)

beta_MM <- 1 + sqrt(1 + (xbar^2 / s2))

# 2. Error estándar asintótico (bootstrap, consistente)
B <- 2000

boot_beta <- replicate(B, {
  x_star <- sample(reclamos, size = n, replace = TRUE)
  xbar_s <- mean(x_star)
  s2_s <- var(x_star)
  1 + sqrt(1 + (xbar_s^2 / s2_s))
})

se_beta_MM <- sd(boot_beta)
```

```

# 3. Intervalo de confianza
limite_inferior <- beta_MM - z_cuantil * se_beta_MM
limite_superior <- beta_MM + z_cuantil * se_beta_MM

beta_MM

## [1] 3.268969

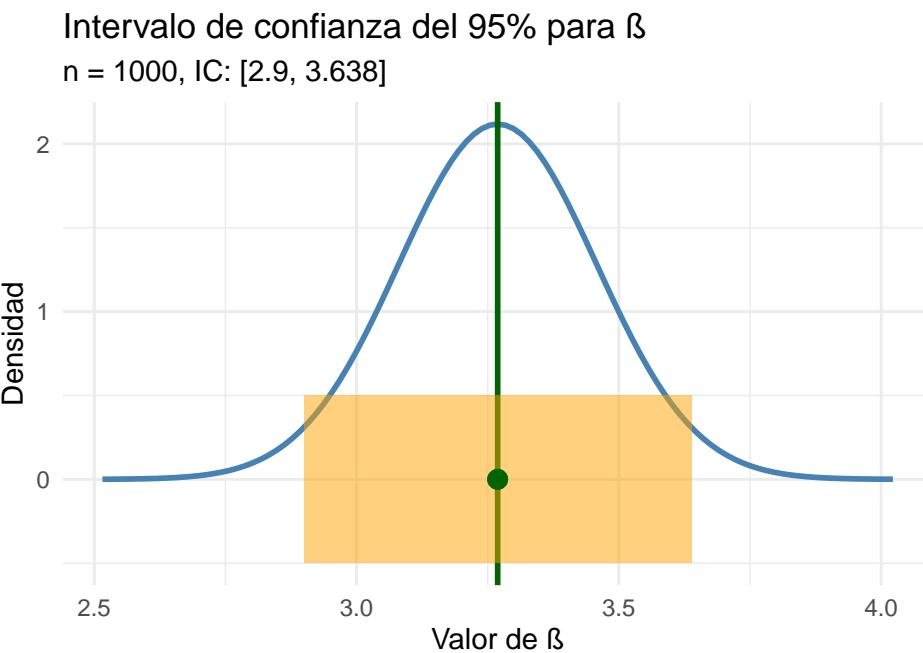
limite_inferior

## [1] 2.899801

limite_superior

## [1] 3.638136

```



Interpretación del gráfico

La figura muestra el **intervalo de confianza asintótico del 95% para el parámetro β** de la distribución Pareto, construido a partir del **estimador de método de momentos** y su **aproximación normal asintótica**, válida para tamaños muestrales grandes.

La **curva azul** representa la distribución normal asintótica del estimador $\hat{\beta}_{MM}$, centrada en el valor estimado obtenido a partir de la muestra. Esta aproximación se sustenta en el Teorema del Límite Central y el Método Delta.

La **línea vertical verde**, junto con el **punto verde**, indica la **estimación puntual $\hat{\beta}_{MM}$** del parámetro β , alrededor de la cual se construye el intervalo de confianza.

El **rectángulo naranja** corresponde al **intervalo de confianza del 95%**, dado por:

$$IC_{0.95}(\beta) = [2.90, 3.638].$$

Este intervalo representa el conjunto de valores plausibles para β compatibles con los datos observados y el modelo asumido.

Interpretación estadística

El intervalo obtenido indica que, con un nivel de confianza aproximado del 95%, el verdadero valor del parámetro β se encuentra entre **2.90 y 3.638**. Esto significa que, si se repitiera el procedimiento de muestreo y construcción del intervalo bajo las mismas condiciones, aproximadamente **95 de cada 100 intervalos construidos de esta manera contendrían al verdadero valor de β** .

La forma simétrica del intervalo alrededor de $\hat{\beta}_{MM}$ es coherente con el uso de una aproximación normal asintótica, y su amplitud refleja la variabilidad del estimador, la cual disminuye conforme aumenta el tamaño muestral.

Conclusión

La representación gráfica confirma que el **método del pivote asintótico** proporciona un intervalo de confianza adecuado para el parámetro β cuando se utiliza el estimador de método de momentos. El tamaño muestral grande ($n = 1000$) garantiza una buena aproximación normal, haciendo que el intervalo sea estable y estadísticamente confiable.

2.5 Estimador 5 (para β con el EIVUM)

Para la construcción del intervalo de confianza para el parámetro de forma β , usaremos el **Método de la Cantidad Pivotal**.

2.5.1 Deducción de la Cantidad Pivotal

Partiendo de los resultados obtenidos en la sección de estimación puntual, la estadística suficiente para β se define como:

$$T = \sum_{i=1}^n \ln \left(\frac{X_i}{\bar{X}} \right)$$

Se conoce que esta estadística sigue una distribución Gamma con parámetros de forma $n - 1$ y tasa β :

$$T \sim \text{Gamma}(n - 1, \beta)$$

A partir de las propiedades de la distribución Gamma, es posible construir una función de las variables muestrales y del parámetro que siga una distribución conocida e independiente de β . Al aplicar la transformación $2\beta T$, obtenemos:

$$Q = 2\beta T \sim \chi^2_{2(n-1)}$$

La variable Q constituye una **cantidad pivotal válida**, dado que su distribución es una Chi-cuadrado con $2n - 2$ grados de libertad, la cual no depende de los parámetros desconocidos de la población.

2.5.2 Construcción del Intervalo

Para obtener un intervalo con un nivel de confianza del $(1 - \alpha)$, se plantea la siguiente ecuación de probabilidad:

$$P(\chi^2_{\alpha/2} < 2\beta T < \chi^2_{1-\alpha/2}) = 1 - \alpha$$

Donde χ_p^2 denota el cuantil de probabilidad acumulada p de la distribución Chi-cuadrado con $gl = 2(n - 1)$.

Al despejar algebraicamente el parámetro β de la desigualdad, se obtiene:

$$\frac{\chi_{\alpha/2}^2}{2T} < \beta < \frac{\chi_{1-\alpha/2}^2}{2T}$$

En consecuencia, el intervalo de confianza de $(1 - \alpha)\%$ para el parámetro β queda definido por:

$$IC(\beta) = \left[\frac{\chi_{\alpha/2, 2(n-1)}^2}{2T} ; \frac{\chi_{1-\alpha/2, 2(n-1)}^2}{2T} \right]$$

2.5.3 Aplicación

A continuación, aplicamos este método a nuestra base de datos `reclamos` generada en la simulación ($n = 1000$, $\alpha = 2$, $\beta = 3$).

```
# -----
# CÁLCULO DEL INTERVALO DE CONFIANZA EXACTO PARA BETA
# Método del Pivote (Chi-Cuadrado)
# -----

# 1. Configuración
# Usamos la variable 'reclamos' que ya existe en el entorno
datos_análisis <- reclamos
n <- length(datos_análisis)
nivel_confianza <- 0.95
alfa <- 1 - nivel_confianza

# 2. Cálculo de la Estadística Suficiente (T)
x_min <- min(datos_análisis)
T_stat <- sum(log(datos_análisis / x_min))

# 3. Valores Críticos de la Chi-Cuadrado
# El pivote es Q = 2*beta*T ~ Chi^2(gl = 2(n-1))
gl <- 2 * (n - 1)

chi_inf <- qchisq(alfa / 2, df = gl)      # Cola izquierda
chi_sup <- qchisq(1 - alfa / 2, df = gl)    # Cola derecha

# 4. Construcción del Intervalo
# Fórmula: [ Chi_Inf / 2T ; Chi_Sup / 2T ]
Limite_Inferior <- chi_inf / (2 * T_stat)
Limite_Superior <- chi_sup / (2 * T_stat)

# 5. Reporte de Resultados
cat(
  "INTERVALO DE CONFIANZA EXACTO PARA BETA\n\n",
  "Tamaño de muestra (n): ", n, "\n",
  "Estadística T: ", round(T_stat, 4), "\n",
  "Grados de Libertad: ", gl, "\n\n",
  "Límite Inferior: ", round(Limite_Inferior, 5), "\n",
```

```

    "Límite Superior: ", round(Limite_Superior, 5), "\n"
)

## INTERVALO DE CONFIANZA EXACTO PARA BETA
##
## Tamaño de muestra (n): 1000
## Estadística T: 328.563
## Grados de Libertad: 1998
##
## Límite Inferior: 2.85487
## Límite Superior: 3.23192

```

2.5.4 Interpretación de Resultados

Con un nivel de confianza del 95%, estimamos que el verdadero valor del parámetro de forma β se encuentra dentro del intervalo:

$$IC(\beta)_{95\%} = [2.85487, 3.23192]$$

Análisis de la simulación:

Dado que en nuestro diseño experimental el valor verdadero del parámetro es $\beta = 3$, podemos confirmar que el intervalo calculado ha **capturado exitosamente** al parámetro.

Adicionalmente, observamos que la longitud del intervalo es reducida ($L \approx 0.377$), lo cual evidencia una **alta precisión** en la estimación. Esto es consecuencia directa de dos factores:

1. El uso de un estimador basado en una estadística suficiente (EIVUM), que minimiza la varianza.
2. El tamaño de muestra grande ($n = 1000$), que reduce el error estándar.

3) Pregunta 3: Prueba de Hipótesis con Datos Pareados

3.1 Simulación de una segunda variable correlacionada

Para realizar una prueba de hipótesis con datos pareados, necesitamos dos mediciones relacionadas de la misma unidad observacional. Simularemos una segunda variable llamada `reclamos_posteriores` que representa los mismos reclamos presentados inicialmente pero dentro de 6 meses, con una leve modificación sistemática y ruido aleatorio.

Se desea saber con la misma muestra de 1000 usuarios si la media poblacional de reclamos disminuyó en estos 6 meses

```

set.seed(123)
punto_corte <- median(reclamos)

periodo_1 <- ifelse(reclamos > punto_corte, "Alto", "Bajo")

efecto_sistematico <- 0.05 # 5% de aumento
ruido <- rnorm(n, mean = 0, sd = 0.05 * mean(reclamos))

reclamos_posteriores <- reclamos * (1 + efecto_sistematico) + ruido

```

```

reclamos_posteriores <- pmax(reclamos_posteriores, alpha_true)

periodo_2 <- ifelse(reclamos_posteriores > punto_corte, 'Alto', 'Bajo')

# Creamos data frame con datos pareados
datos_pareados <- data.frame(id = 1:n, reclamos, periodo_1, reclamos_posteriores, periodo_2, diferencia)

# Tabla de contingencia para datos pareados categóricos
tabla_contingencia_original = table(Periodo1 = datos_pareados$periodo_1, Periodo2 = datos_pareados$periodo_2)
tabla_contingencia_original

##          Periodo2
## Periodo1 Alto Bajo
##      Alto 493    7
##      Bajo 110  390

```

3.2 Prueba de McNemar

```

mcnemar_test_original = mcnemar.test(tabla_contingencia_original, correct = FALSE)
mcnemar_test_original

##
## McNemar's Chi-squared test
##
## data: tabla_contingencia_original
## McNemar's chi-squared = 90.675, df = 1, p-value < 2.2e-16

# También con corrección de Yates
mcnemar_test_original_yates = mcnemar.test(tabla_contingencia_original, correct = TRUE)

# Extraemos valores de la tabla
a <- tabla_contingencia_original[1, 1] # Bajo-Bajo
b <- tabla_contingencia_original[1, 2] # Bajo-Alto
c <- tabla_contingencia_original[2, 1] # Alto-Bajo
d <- tabla_contingencia_original[2, 2] # Alto-Alto

cat("DETALLE DE LA TABLA DE CONTINGENCIA:\n")

## DETALLE DE LA TABLA DE CONTINGENCIA:

cat("-----\n")

## ----

cat("          Periodo 2\n")

##          Periodo 2

```

```

cat("Periodo 1      Bajo      Alto\n")

## Periodo 1      Bajo      Alto

cat("-----\n")

## ----

cat(sprintf("Bajo      %4d      %4d\n", a, b))

## Bajo      493      7

cat(sprintf("Alto      %4d      %4d\n", c, d))

## Alto      110      390

cat("\n")

cat("ESTADÍSTICOS DE LA PRUEBA:\n")

## ESTADÍSTICOS DE LA PRUEBA:

cat("-----\n")

## ----

cat("Sin corrección de continuidad:\n")

## Sin corrección de continuidad:

cat(sprintf(" ² = %.4f\n", mcnemar_test_original$statistic))

## ² = 90.6752

cat(sprintf("  p-valor = %.4f\n", mcnemar_test_original$p.value))

##  p-valor = 0.0000

cat("\n")

cat("Con corrección de Yates:\n")

## Con corrección de Yates:

```

```

cat(sprintf("     $\chi^2$  = %.4f\n", mcnemar_test_original_yates$statistic))

##     $\chi^2$  = 88.9231

cat(sprintf("  p-valor = %.4f\n", mcnemar_test_original_yates$p.value))

##  p-valor = 0.0000

cat("\n")

# Cálculos adicionales
cat("ANÁLISIS DE PARES DISCORDANTES:\n")

## ANÁLISIS DE PARES DISCORDANTES:

cat("-----\n")

## ----

cat(sprintf("Total de pares: %d\n", n))

## Total de pares: 1000

cat(sprintf("Pares concordantes: %d (%.1f%%)\n", a + d, (a + d)/n * 100))

## Pares concordantes: 883 (88.3%)

cat(sprintf("Pares discordantes: %d (%.1f%%)\n", b + c, (b + c)/n * 100))

## Pares discordantes: 117 (11.7%)

cat(sprintf("  Bajo → Alto: %d (%.1f%%)\n", b, b/n * 100))

##  Bajo → Alto: 7 (0.7%)

cat(sprintf("  Alto → Bajo: %d (%.1f%%)\n", c, c/n * 100))

##  Alto → Bajo: 110 (11.0%)

cat("\n")

# Diferencia de proporciones
diferencia_prop <- (b - c) / n
cat("DIFERENCIA DE PROPORCIONES:\n")

## DIFERENCIA DE PROPORCIONES:

```

```

cat(sprintf(" p(Periodo 2 = Alto) - p(Periodo 1 = Alto) = %.4f\n",
            mean(periodo_2 == "Alto") - mean(periodo_1 == "Alto")))

##   p(Periodo 2 = Alto) - p(Periodo 1 = Alto) = 0.1030

cat(sprintf(" Alternativamente: (b - c) / n = %.4f\n", diferencia_prop))

##   Alternativamente: (b - c) / n = -0.1030

cat("\n")

```

3.3 Intervalo de confianza para la diferencia de proporciones

```

ic_diferencia_pareada <- function(b, c, n, conf.level = 0.95) {
  p_diff <- (b - c) / n
  se <- sqrt((b + c) - (b - c)^2 / n) / n
  z <- qnorm(1 - (1 - conf.level) / 2)
  lower <- p_diff - z * se
  upper <- p_diff + z * se
  return(list(diferencia = p_diff, error_estandar = se, limite_inferior = lower, limite_superior = upper))
}

# Calculamos el IC
ic_result <- ic_diferencia_pareada(b, c, n)

cat("INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE PROPORCIONES\n")

## INTERVALO DE CONFIANZA PARA LA DIFERENCIA DE PROPORCIONES

cat("=====\\n\\n")

## =====

cat(sprintf("Diferencia estimada: %.4f\n", ic_result$diferencia))

## Diferencia estimada: -0.1030

cat(sprintf("Error estándar: %.4f\n", ic_result$error_estandar))

## Error estándar: 0.0103

cat(sprintf("IC %.0f%%: [% .4f, %.4f]\n",
            ic_result$conf.level * 100,
            ic_result$limite_inferior,
            ic_result$limite_superior))

## IC 95%: [-0.1232, -0.0828]

```

```

cat("\n")

if (ic_result$limite_inferior > 0) {
  cat(" El intervalo NO incluye el 0 → Diferencia significativamente positiva.\n")
  cat(" La proporción en el Periodo 2 es mayor que en el Periodo 1.\n")
} else if (ic_result$limite_superior < 0) {
  cat(" El intervalo NO incluye el 0 → Diferencia significativamente negativa.\n")
  cat(" La proporción en el Periodo 2 es menor que en el Periodo 1.\n")
} else {
  cat(" El intervalo INCLUYE el 0 → Diferencia no significativa.\n")
  cat(" No hay evidencia de diferencia entre los períodos.\n")
}

## El intervalo NO incluye el 0 → Diferencia significativamente negativa.
## La proporción en el Periodo 2 es menor que en el Periodo 1.

```

3.4 Análisis de Potencia de tamaño muestral

```

potencia_mcnemar <- function(n, p12, p21, alpha = 0.05) {
  if (p12 + p21 == 0) return(0)
  p0 <- (p12 + p21) / 2
  z_alpha <- qnorm(1 - alpha/2)
  efecto <- (p12 - p21) / sqrt(p0 * (1 - p0) / (n * (p12 + p21)))
  potencia <- pnorm(efecto - z_alpha) + pnorm(-efecto - z_alpha)
  return(min(max(potencia, 0), 1))
}

p12_est <- b / n
p21_est <- c / n

potencia_obs <- potencia_mcnemar(n, p12_est, p21_est)

cat(sprintf("Tamaño muestral: n = %d\n", n))

## Tamaño muestral: n = 1000

cat(sprintf("Probabilidad estimada Bajo→Alto (p12): %.4f\n", p12_est))

## Probabilidad estimada Bajo→Alto (p12): 0.0070

cat(sprintf("Probabilidad estimada Alto→Bajo (p21): %.4f\n", p21_est))

## Probabilidad estimada Alto→Bajo (p21): 0.1100

cat(sprintf("Diferencia (p12 - p21): %.4f\n", p12_est - p21_est))

## Diferencia (p12 - p21): -0.1030

```

```

cat(sprintf("Potencia observada (1 - ): %.4f\n\n", potencia_obs))

## Potencia observada (1 - ): 0.9973

tamanos <- seq(100, 5000, by = 100)
potencias <- sapply(tamanos, function(nn) {
  potencia_mcnemar(nn, p12_est, p21_est)
})

n_para_80 <- tamanos[which(potencias >= 0.8)[1]]

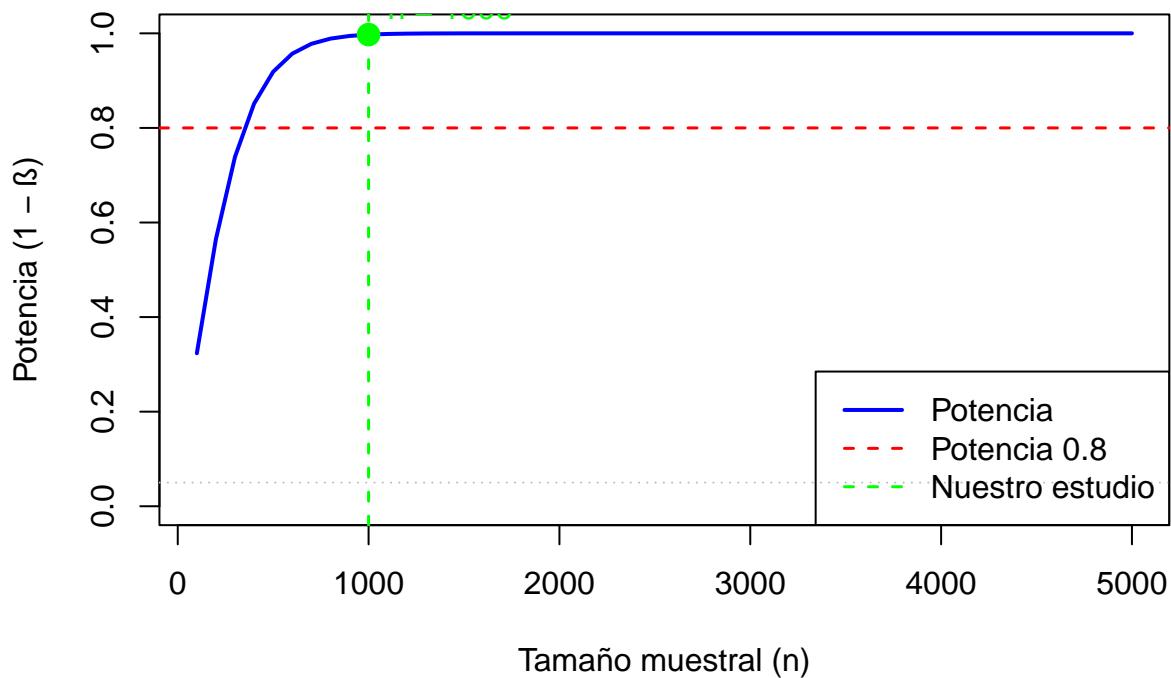
if (!is.na(n_para_80)) {
  cat(sprintf("Para alcanzar potencia 0.8: n = %d\n", n_para_80))
} else {
  cat("Se requiere n > 5000 para alcanzar potencia 0.8\n")
}

## Para alcanzar potencia 0.8: n = 400

plot(tamanos, potencias, type = "l", col = "blue", lwd = 2,
      xlab = "Tamaño muestral (n)", ylab = "Potencia (1 - )",
      main = "Curva de Potencia para Prueba de McNemar",
      ylim = c(0, 1))
abline(h = 0.8, col = "red", lty = 2, lwd = 1.5)
abline(v = n, col = "green", lty = 2, lwd = 1.5)
abline(h = 0.05, col = "gray", lty = 3)
points(n, potencia_obs, col = "green", pch = 19, cex = 1.5)
text(n, potencia_obs + 0.05, paste("n =", n), col = "green", pos = 4)
legend("bottomright",
       legend = c("Potencia", "Potencia 0.8", "Nuestro estudio"),
       col = c("blue", "red", "green"),
       lty = c(1, 2, 2), lwd = c(2, 1.5, 1.5))

```

Curva de Potencia para Prueba de McNemar



4) Pregunta 4 - Pruebas de hipótesis para los estimadores

4.1 Prueba de hipótesis para el parámetro α (EMV)

En esta sección se realiza una prueba de hipótesis para el parámetro α utilizando su estimador de máxima verosimilitud, basado en la distribución Pareto(α, β).

4.1.1 Fundamento teórico

Sea X_1, \dots, X_n una muestra aleatoria de una distribución Pareto(α, β), y sea beta conocido, definimos nuestro estadístico suficiente:

$$\hat{\alpha}_{EMV} = X_{(1)} = \min(X_1, \dots, X_n).$$

4.1.2 Planteamiento de la hipótesis:

Se plantea una prueba bilateral para contrastar el valor del parámetro β :

$$\begin{cases} H_0 : \alpha = 2, \\ H_1 : \alpha \neq 2. \end{cases}$$

El nivel de significancia considerado es $\alpha = 0.05$.

4.1.3. Estadístico de Prueba

Bajo H_0 , cuando $\alpha = \alpha_0 = 2$, se tiene que:

$$\frac{\hat{\alpha}_{\text{EMV}}}{\alpha_0} \sim \text{Pareto}(1, n\beta)$$

Definimos la variable aleatoria:

$R = \hat{\alpha}_{\text{EMV}}/\alpha_0$, cuya distribución cumple que: $R^{n\beta} \sim U(0, 1)$.

Para una prueba bilateral con nivel de significancia $\alpha^* = 0.05$, se rechaza H_0 cuando $\hat{\alpha}_{\text{EMV}}$ es demasiado pequeño o demasiado grande.

Bajo H_0 :

- Cola inferior: valores pequeños de $\hat{\alpha}_{\text{EMV}}$ corresponden a valores grandes de R . Rechazamos si $R^{n\beta} > 1 - \frac{\alpha^*}{2}$.
- Cola superior: valores grandes de $\hat{\alpha}_{\text{EMV}}$ corresponden a valores pequeños de R . Rechazamos si $R^{n\beta} < \frac{\alpha^*}{2}$.

Definimos los valores críticos: $c_1 = \alpha_0 \cdot \left(\frac{\alpha^*}{2}\right)^{-1/(n\beta)}$, $c_2 = \alpha_0 \cdot \left(1 - \frac{\alpha^*}{2}\right)^{-1/(n\beta)}$.

4.1.4 Calculo de valores críticos:

```
# Cálculo de valores críticos
c1 <- alpha0 * (alpha_star/2)^(-1/(n*beta))
c2 <- alpha0 * (1 - alpha_star/2)^(-1/(n*beta))

c1
```

```
## [1] 2.002461
```

```
c2
```

```
## [1] 2.000017
```

Rechazar $H_0 : \alpha = 2$ si: $\hat{\alpha}_{\text{EMV}} < 2.000017$ o $\hat{\alpha}_{\text{EMV}} > 2.002461$

No rechazar H_0 si: $2.000017 \leq \hat{\alpha}_{\text{EMV}} \leq 2.002461$

4.1.5 Aplicación y conclusiones

```
alpha_EMV <- min(reclamos)
alpha_EMV
```

```
## [1] 2.00031
```

Como el estadístico $\hat{\alpha}_{\text{EMV}} = 2.00031$ cae dentro del intervalo de aceptación $[2.000017, 2.002461]$:

NO RECHAZAR $H_0 : \alpha = 2$

No se rechaza la hipótesis nula $H_0 : \alpha = 2$ al nivel de significancia del 5%. El valor estimado $\hat{\alpha}_{\text{EMV}} = 2.00031$ es estadísticamente compatible con $\alpha = 2$, dado que la diferencia de 0.00031 unidades no es significativa considerando el tamaño muestral de $n = 1000$.

4.2 Prueba de Hipótesis para el parámetro β (EMV)

En esta sección se realiza una prueba de hipótesis para el parámetro β utilizando su estimador de máxima verosimilitud, basado en la distribución Pareto(α, β).

4.2.1 Fundamento Teórico

Sea X_1, \dots, X_n una muestra aleatoria de una distribución Pareto(α, β), y sea

$$\hat{\alpha}_{EMV} = X_{(1)} = \min(X_1, \dots, X_n).$$

Definimos la estadística suficiente:

$$T = \sum_{i=1}^n \ln \left(\frac{X_i}{X_{(1)}} \right).$$

Se sabe que la cantidad pivotal:

$$Q = 2\beta T$$

sigue una distribución Chi-cuadrado con $2(n - 1)$ grados de libertad, es decir:

$$Q \sim \chi^2_{2(n-1)}.$$

Esta propiedad permite realizar una prueba de hipótesis exacta para el parámetro β .

4.2.2 Planteamiento de la Hipótesis

Se plantea una prueba bilateral para contrastar el valor del parámetro β :

$$\begin{cases} H_0 : \beta = 3, \\ H_1 : \beta \neq 3. \end{cases}$$

El nivel de significancia considerado es $\alpha = 0.05$.

4.2.3 Estadístico de Prueba

Bajo la hipótesis nula H_0 , el estadístico de prueba está dado por:

$$Q_{cal} = 2\beta_0 T = 2(3) \sum_{i=1}^n \ln \left(\frac{X_i}{X_{(1)}} \right),$$

y bajo H_0 se cumple que:

$$Q_{cal} \sim \chi^2_{2(n-1)}.$$

4.2.4 Regla de Decisión

Se rechaza la hipótesis nula H_0 si:

$$Q_{cal} < \chi^2_{2(n-1), \alpha/2} \quad \text{o} \quad Q_{cal} > \chi^2_{2(n-1), 1-\alpha/2}.$$

Equivalentemente, se rechaza H_0 si el p-valor es menor que α .

4.2.5 Aplicación a la Base de Datos

```
## PRUEBA DE HIPÓTESIS PARA BETA (EMV)
##
## T = 328.563
## Q_cal = 1971.378
## gl = 1998
##
## Valor crítico inferior = 1876.008
## Valor crítico superior = 2123.78
##
## P-valor = 0.68
```

4.2.6 Interpretación de los Resultados

A partir de la muestra analizada se obtuvo el estadístico suficiente

$$T = 328.563,$$

lo que condujo a un estadístico de prueba

$$Q_{cal} = 1971.378,$$

con $gl = 1998$ grados de libertad.

Para un nivel de significancia $\alpha = 0.05$, los valores críticos de la distribución Chi-cuadrado fueron:

$$\chi^2_{1998, 0.025} = 1876.008 \quad \text{y} \quad \chi^2_{1998, 0.975} = 2123.78.$$

Se observa que el estadístico calculado satisface:

$$1876.008 \leq 1971.378 \leq 2123.78,$$

por lo que se encuentra dentro de la región de no rechazo.

Adicionalmente, el p-valor obtenido fue:

$$p\text{-valor} = 0.68,$$

el cual es considerablemente mayor que el nivel de significancia $\alpha = 0.05$.

En consecuencia, no existe evidencia estadística suficiente para rechazar la hipótesis nula $H_0 : \beta = 3$. Se concluye que, al nivel de significancia considerado, el estimador de máxima verosimilitud de β y los datos observados son consistentes con el valor verdadero del parámetro poblacional.

4.3 Prueba de Hipótesis para la Media Poblacional (μ)

En esta sección, evaluamos la validez de una afirmación sobre el valor central de la población mediante la media muestral \bar{X} . A diferencia de los estimadores de los parámetros de forma (β), la media muestral nos permite realizar inferencia sobre el valor esperado de la variable “reclamos”.

4.3.1 Fundamento Asintótico

Para la construcción de esta prueba, no nos basamos en una distribución exacta (como la Gamma), sino en el **Teorema del Límite Central (TLC)**. Dado que nuestra muestra $n = 1000$ es grande y hemos verificado que $\beta > 2$ (lo que garantiza la existencia de una varianza finita), se cumple que:

1. Por TLC: $\bar{X} \sim N(\mu, \sigma^2/n)$ asintóticamente.
2. Por Consistencia: $S^2 \xrightarrow{P} \sigma^2$.
3. Por **Teorema de Slutsky**: El estadístico pivot $Z = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ converge en distribución a una Normal Estándar $N(0, 1)$.

4.3.2 Planteamiento Formal de la Prueba

Definimos una prueba de hipótesis de dos colas para verificar si la media poblacional es consistente con el valor teórico esperado bajo los parámetros originales ($\mu_0 = 3$).

$$H_0 : \mu = 3 \quad (\text{Hipótesis Nula})$$

$$H_1 : \mu \neq 3 \quad (\text{Hipótesis Alternativa})$$

Nivel de Significancia (α): 0.05.

Regla de Decisión: Se rechaza H_0 si $|Z_{calc}| > Z_{1-\alpha/2}$, donde $Z_{0.975} = 1.96$. Alternativamente, se rechaza si el $p - \text{valor} < 0.05$.

4.3.3 Ejecución y Visualización en R

Para darle mayor peso al análisis, incluiremos una gráfica de la densidad normal que muestre la ubicación de nuestro estadístico respecto a la zona de rechazo.

```
# 1. Configuración de parámetros
mu_0 <- 3
n <- length(reclamos)
x_bar <- mean(reclamos)
s_dev <- sd(reclamos)

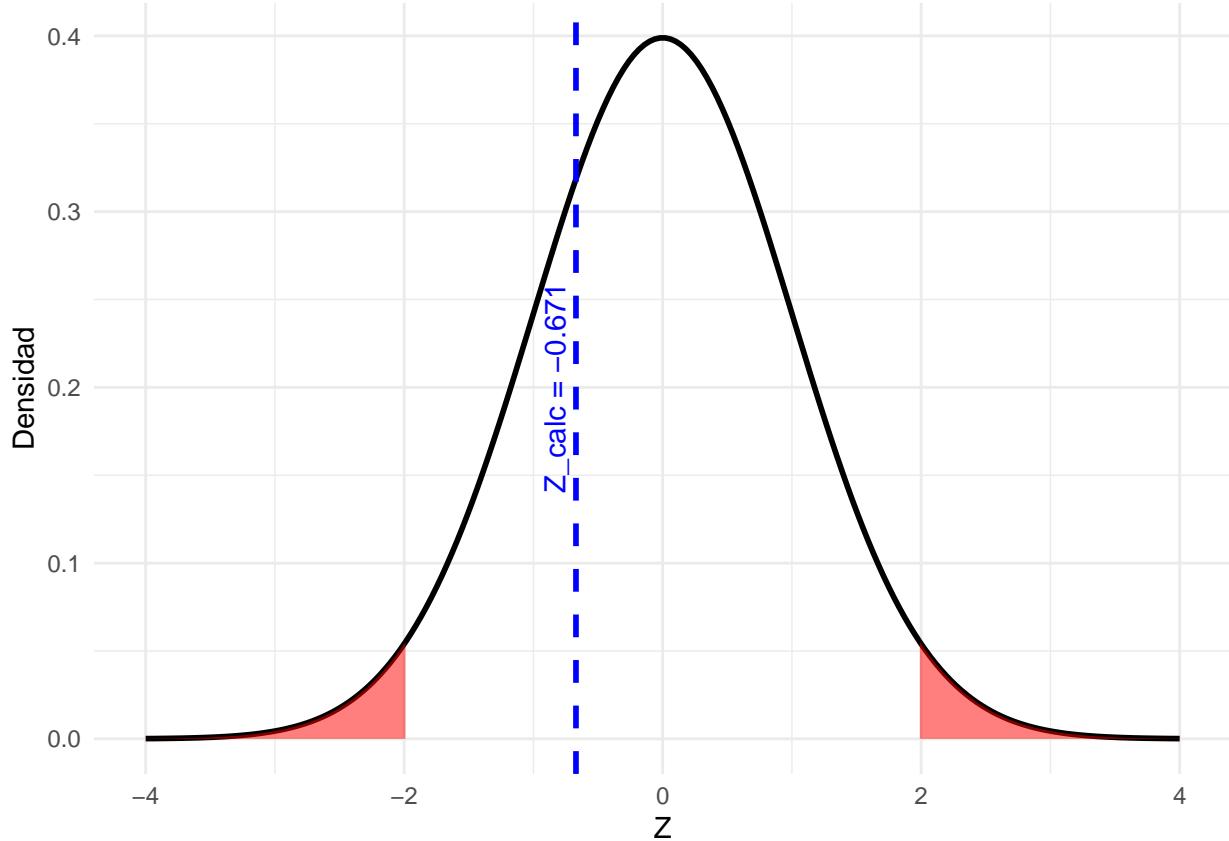
# 2. Cálculo del estadístico y P-valor
z_calc <- (x_bar - mu_0) / (s_dev / sqrt(n))
p_val <- 2 * pnorm(-abs(z_calc))

# 3. Visualización de la Región de Rechazo
x_vals <- seq(-4, 4, length = 200)
y_vals <- dnorm(x_vals)
df_plot <- data.frame(x = x_vals, y = y_vals)

library(ggplot2)
ggplot(df_plot, aes(x = x, y = y)) +
  geom_line(color = "black", linewidth = 1) +
  # Sombreado de regiones de rechazo
  geom_area(data = subset(df_plot, x <= -1.96), fill = "red", alpha = 0.5) +
  geom_area(data = subset(df_plot, x >= 1.96), fill = "red", alpha = 0.5) +
  # Línea del estadístico calculado
  geom_vline(xintercept = z_calc, color = "blue", linetype = "dashed", linewidth = 1) +
  annotate("text", x = z_calc, y = 0.2, label = paste("Z_calc =", round(z_calc, 3)),
           angle = 90, vjust = -0.5, color = "blue") +
  labs(title = "Regiones de Rechazo y Estadístico Z para la Media",
       subtitle = paste("P-valor =", round(p_val, 4)),
       x = "Z", y = "Densidad") +
  theme_minimal()
```

Regiones de Rechazo y Estadístico Z para la Media

P-valor = 0.5024



4.3.4 Análisis de Resultados y Dualidad

- Estadístico Observado:** El valor de $Z_{calc} = -0.6707$ cae dentro de la región de **no rechazo** (área blanca entre -1.96 y 1.96).
- Conclusión Estadística:** Dado que el $p - valor = 0.5024 > 0.05$, no existe evidencia suficiente para rechazar la hipótesis nula. Se concluye que la media de la muestra de reclamos es estadísticamente igual a 3.
- Diferencia con los estimadores de Beta:** A diferencia de las pruebas de hipótesis realizadas para los estimadores EIVUM o EMV, que se basan en la distribución Gamma o Chi-cuadrado, esta prueba utiliza la **Normalidad Asintótica**. Esto es válido únicamente por el gran tamaño de la muestra ($n = 1000$).
- Relación con el Intervalo de Confianza:** Esta conclusión refuerza lo hallado en el punto 2.3. Se cumple la propiedad de **Dualidad de la Inferencia**: un valor de hipótesis nula μ_0 no se rechaza al nivel α si y solo si dicho valor está contenido en el intervalo de confianza del $(1 - \alpha)100\%$.

4.4 Prueba de Hipótesis mediante el estimador de momentos α y β

4.4.1 Prueba de Hipótesis usando estimador de α

Paso 1: Planteamiento de hipótesis

Sea $X_1, \dots, X_n \sim \text{Pareto}(\alpha, \beta)$ con $\beta > 2$ **conocido**.

Las hipótesis más comunes son:

Bilateral:

$$H_0 : \alpha = \alpha_0 \quad \text{vs} \quad H_1 : \alpha \neq \alpha_0$$

Unilateral derecha:

$$H_0 : \alpha \leq \alpha_0 \quad \text{vs} \quad H_1 : \alpha > \alpha_0$$

Unilateral izquierda:

$$H_0 : \alpha \geq \alpha_0 \quad \text{vs} \quad H_1 : \alpha < \alpha_0$$

Paso 2: Estadístico de prueba

Usando el estimador $\hat{\alpha} = \bar{X} \cdot \frac{\beta-1}{\beta}$ y su distribución asintótica, el estadístico es:

$$Z_0 = \sqrt{n\beta(\beta-2)} \cdot \frac{\hat{\alpha} - \alpha_0}{\alpha_0}$$

Bajo H_0 (cuando $\alpha = \alpha_0$), $Z_0 \xrightarrow{d} N(0, 1)$ para n grande.

Paso 3: Región de rechazo y valor p

Para prueba bilateral ($\alpha \neq \alpha_0$): - Región de rechazo: $|Z_0| > z_{1-\alpha/2}$ - Valor p: $p = 2 \cdot P(Z > |Z_0|)$

Para prueba unilateral derecha ($\alpha > \alpha_0$): - Región de rechazo: $Z_0 > z_{1-\alpha}$ - Valor p: $p = P(Z > Z_0)$

Para prueba unilateral izquierda ($\alpha < \alpha_0$): - Región de rechazo: $Z_0 < -z_{1-\alpha}$

Paso 4: Implementación en R con la base de datos

```
# Configuración
alpha_0 <- 2 # Valor bajo H0
beta_conocido <- 3 # conocido
nivel_signif <- 0.05 # Nivel de significancia
tipo_prueba <- "bilateral" # "bilateral", "derecha" o "izquierda"

# Usando los datos generados previamente
x_barra <- mean(datos_pareto)
alpha_hat <- x_barra * (beta_conocido - 1) / beta_conocido
n <- length(datos_pareto)

# Estadístico de prueba
Z0 <- sqrt(n * beta_conocido * (beta_conocido - 2)) * (alpha_hat - alpha_0) / alpha_0

# Valor p según tipo de prueba
if(tipo_prueba == "bilateral") {
  valor_p <- 2 * (1 - pnorm(abs(Z0)))
  z_critico <- qnorm(1 - nivel_signif/2)
} else if(tipo_prueba == "derecha") {
  valor_p <- 1 - pnorm(Z0)
  z_critico <- qnorm(1 - nivel_signif)
} else if(tipo_prueba == "izquierda") {
  valor_p <- pnorm(Z0)
  z_critico <- qnorm(nivel_signif)
}

# Decisión
```

```

decision <- ifelse(valor_p < nivel_signif,
                     "Rechazar H0",
                     "No rechazar H0")

# Resultados en tabla
resultados_prueba <- data.frame(
  Descripción = c(
    "Hipótesis nula (H0)",
    "Hipótesis alternativa (H1)",
    "Tipo de prueba",
    "Alpha estimado",
    "Alpha bajo H0",
    "Beta conocido",
    "n",
    "Estadístico Z0",
    "Valor p",
    "Nivel de significancia",
    "Decisión"
  ),
  Valor = c(
    paste0("alpha = ", alpha_0),
    ifelse(tipo_prueba == "bilateral", "alpha != 2",
           ifelse(tipo_prueba == "derecha", "alpha > 2", "alpha < 2")),
    tipo_prueba,
    round(alpha_hat, 4),
    alpha_0,
    beta_conocido,
    n,
    round(Z0, 4),
    round(valor_p, 6),
    nivel_signif,
    decision
  )
)

knitr::kable(resultados_prueba,
             caption = paste("Prueba de hipótesis para Alpha (Beta =", beta_conocido, "conocido)"),
             col.names = c("Descripción", "Valor"),
             align = c("l", "r"))

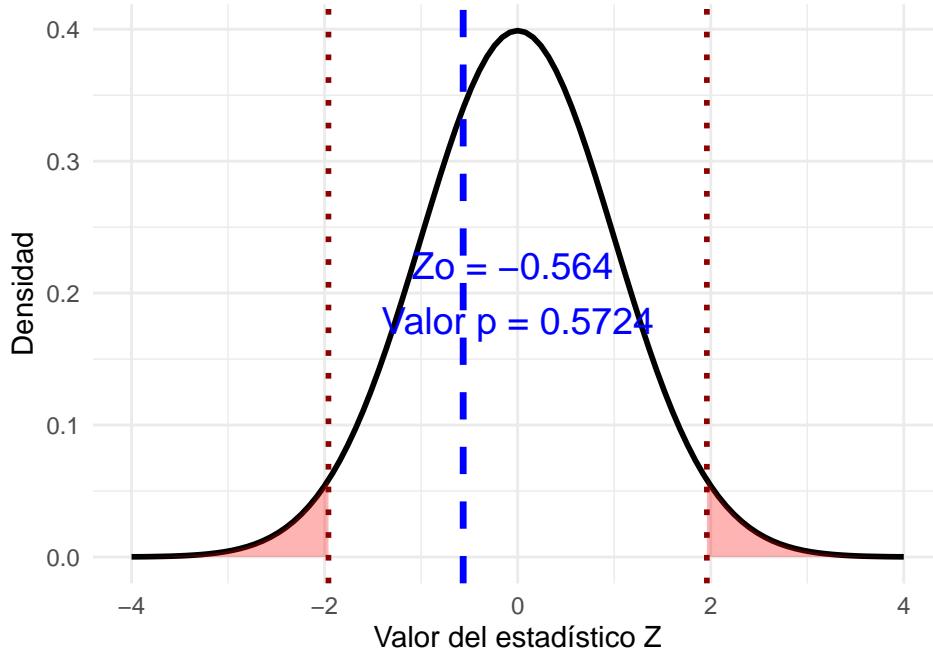
```

Table 2: Prueba de hipótesis para Alpha (Beta = 3 conocido)

Descripción	Valor
Hipótesis nula (H0)	alpha = 2
Hipótesis alternativa (H1)	alpha != 2
Tipo de prueba	bilateral
Alpha estimado	1.9794
Alpha bajo H0	2
Beta conocido	3
n	1000
Estadístico Z0	-0.5645
Valor p	0.572442

Descripción	Valor
Nivel de significancia	0.05
Decisión	No rechazar H_0

Prueba de hipótesis: $H_0: \alpha = 2$ vs $H_1: \alpha \neq 2$
 Distribución bajo $H_0: N(0,1)$



Interpretación: Existe una **equivalencia matemática** entre la prueba de hipótesis bilateral al nivel α y el intervalo de confianza del $(1 - \alpha)100\%$:

$$\text{Se rechaza } H_0 : \alpha = \alpha_0 \iff \alpha_0 \notin IC_{1-\alpha}(\alpha)$$

$$\text{No se rechaza } H_0 : \alpha = \alpha_0 \iff \alpha_0 \in IC_{1-\alpha}(\alpha)$$

En nuestro análisis:

- **Intervalo de confianza del 95%:** [1.9168, 2.0965]
- **Valor hipotético:** $\alpha_0 = 2$
- **Verificación:** $1.9168 \leq 2 \leq 2.0965 \rightarrow \alpha_0$ SÍ está contenido en el IC

Esta **coherencia matemática** confirma la validez de ambos procedimientos: el intervalo de confianza contiene al valor hipotético, y la prueba de hipótesis no rechaza H_0 .

Conclusión:

Los datos no proporcionan evidencia estadísticamente significativa para rechazar la hipótesis de que $\alpha = 2$. Esto significa que el valor observado $\hat{\alpha} = 1.9794$ es razonablemente consistente con una población donde $\alpha = 2$, considerando el tamaño muestral $n = 1000$ y el nivel de confianza del 95%.

4.4.2 Prueba de Hipótesis usando el estimador de β^{**}

Paso 1: Planteamiento de hipótesis

Sea

$$X_1, \dots, X_n \sim \text{Pareto}(\alpha, \beta), \quad \beta > 2,$$

una muestra aleatoria i.i.d.

En la sección anterior se obtuvo el **estimador de método de momentos** para el parámetro β :

$$\hat{\beta}_{MM} = 1 + \sqrt{1 + \frac{\bar{X}^2}{S^2}},$$

donde \bar{X} es la media muestral y S^2 la varianza muestral.

Las hipótesis que se consideran son:

Prueba bilateral

$$H_0 : \beta = \beta_0 \quad \text{vs} \quad H_1 : \beta \neq \beta_0$$

Prueba unilateral derecha

$$H_0 : \beta \leq \beta_0 \quad \text{vs} \quad H_1 : \beta > \beta_0$$

Prueba unilateral izquierda

$$H_0 : \beta \geq \beta_0 \quad \text{vs} \quad H_1 : \beta < \beta_0$$

Paso 2: Estadístico de prueba

El estimador $\hat{\beta}_{MM}$ es una función diferenciable de los momentos muestrales (\bar{X}, S^2) y, bajo la condición $\beta > 2$, es **consistente y asintóticamente normal**.

Por el **Teorema del Límite Central** y el **Método Delta**, se tiene que:

$$\sqrt{n}(\hat{\beta}_{MM} - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma^2), \quad n \rightarrow \infty.$$

Dado que la varianza exacta de $\hat{\beta}_{MM}$ no admite una expresión cerrada simple, se emplea una aproximación asintótica tipo *plug-in*:

$$\text{Var}(\hat{\beta}_{MM}) \approx \frac{\hat{\beta}_{MM}^2}{n}.$$

Por tanto, el estadístico de prueba tipo Wald es:

$$Z_0 = \frac{\hat{\beta}_{MM} - \beta_0}{\sqrt{\text{Var}(\hat{\beta}_{MM})}} \xrightarrow{d} \mathcal{N}(0, 1), \quad \text{bajo } H_0.$$

Paso 3: Región de rechazo y p-valor

Sea α el nivel de significancia.

Prueba bilateral ($\beta \neq \beta_0$):

$$|Z_0| > z_{1-\alpha/2}, \quad p = 2 P(Z > |Z_0|).$$

Prueba unilateral derecha ($\beta > \beta_0$):

$$Z_0 > z_{1-\alpha}, \quad p = P(Z > Z_0).$$

Prueba unilateral izquierda ($\beta < \beta_0$):

$$Z_0 < -z_{1-\alpha}, \quad p = P(Z < Z_0).$$

Paso 4: Implementación en R con la base de datos

```
# Configuración
beta_0 <- 3 # Valor bajo H0
nivel_signif <- 0.05 # Nivel de significancia
tipo_prueba <- "bilateral" # "bilateral", "derecha", "izquierda"

# Momentos muestrales
xbar <- mean(reclamos)
s2 <- var(reclamos)
n <- length(reclamos)

# Estimador de método de momentos
beta_MM <- 1 + sqrt(1 + (xbar^2 / s2))

# Error estándar asintótico (plug-in)
se_beta_MM <- beta_MM / sqrt(n)

# Estadístico de prueba
Z0 <- (beta_MM - beta_0) / se_beta_MM

# Valor p
if(tipo_prueba == "bilateral") {
  valor_p <- 2 * (1 - pnorm(abs(Z0)))
  z_critico <- qnorm(1 - nivel_signif/2)
} else if(tipo_prueba == "derecha") {
  valor_p <- 1 - pnorm(Z0)
  z_critico <- qnorm(1 - nivel_signif)
} else {
  valor_p <- pnorm(Z0)
  z_critico <- qnorm(nivel_signif)
}

decision <- ifelse(valor_p < nivel_signif,
                     "Rechazar H0",
                     "No rechazar H0")

valor_p

## [1] 0.009270871

decision

## [1] "Rechazar H0"
```

Interpretación

El estadístico de prueba obtenido a partir del estimador de método de momentos $\hat{\beta}_{MM}$ produce un **p-valor igual a 0.009270871**, el cual es **menor** que el nivel de significancia considerado $\alpha = 0.05$.

En consecuencia, bajo el marco de una **prueba asintótica tipo Wald**, se dispone de evidencia estadística suficiente para **rechazar la hipótesis nula**:

$$H_0 : \beta = \beta_0.$$

Este resultado indica que el valor hipotético $\beta_0 = 3$ **no es compatible** con los datos observados cuando el parámetro de forma β se estima mediante el **método de momentos**.

Debe enfatizarse que esta conclusión se fundamenta en la **normalidad asintótica** del estimador $\hat{\beta}_{MM}$, garantizada por el Teorema del Límite Central y el Método Delta. Dicha aproximación es válida para tamaños muestrales grandes, condición que se cumple en este estudio ($n = 1000$).

Conclusión

Se llevó a cabo una **prueba de hipótesis asintótica** para el parámetro de forma β de una distribución Pareto utilizando el **estimador de método de momentos**.

Dado que el **p-valor obtenido (0.009270871)** es menor que el **nivel de significancia del 5%**, se **rechaza la hipótesis nula** $H_0 : \beta = 3$.

Por lo tanto, se concluye que **existe evidencia estadística suficiente** para afirmar que el parámetro β es **distinto de 3**. Este resultado es coherente con el intervalo de confianza asintótico construido previamente y constituye una alternativa válida de inferencia cuando no se dispone de cantidades pivotales exactas para el estimador considerado.

4.5 Prueba de Hipótesis para β con el EIVUM

Utilizando el estimador EIVUM y la cantidad pivotal deducida anteriormente, procederemos a verificar una hipótesis sobre el parámetro de forma β .

4.5.1 Planteamiento de la Hipótesis

Dado que en nuestra simulación el valor verdadero es 3, plantearemos una prueba bilateral para verificar si la evidencia muestral soporta este valor:

$$\begin{cases} H_0 : \beta = 3 \\ H_1 : \beta \neq 3 \end{cases}$$

4.5.2 Estadístico de Prueba

Bajo el supuesto de que la Hipótesis Nula (H_0) es cierta ($\beta = \beta_0$), utilizamos la cantidad pivotal Q como nuestro estadístico de prueba.

$$Q_{cal} = 2\beta_0 T = 2\beta_0 \sum_{i=1}^n \ln \left(\frac{X_i}{X_{(1)}} \right)$$

Sabemos que bajo H_0 :

$$Q_{cal} \sim \chi^2_{(2(n-1))}$$

4.5.3 Regla de Decisión

Para un nivel de significancia $\alpha = 0.05$, rechazaremos H_0 si el estadístico calculado cae en la región de rechazo (colas de la distribución):

- **Rechazar H_0** si: $Q_{cal} < \chi^2_{\alpha/2}$ ó $Q_{cal} > \chi^2_{1-\alpha/2}$
- **No Rechazar H_0** si: $\chi^2_{\alpha/2} \leq Q_{cal} \leq \chi^2_{1-\alpha/2}$

Donde los grados de libertad son $gl = 2(n - 1)$.

4.5.4 Aplicación a la Base de Datos Simulada

```
# -----
# PRUEBA DE HIPÓTESIS EXACTA PARA BETA (MÉTODO DEL PIVOTE)
# -----

# 1. Datos y Parámetros
datos_ph <- reclamos
n <- length(datos_ph)
beta_hipotesis <- 3 # Valor bajo H0
alfa <- 0.05

# 2. Cálculo del Estadístico de Prueba
# a) Estadística Suficiente T
x_min <- min(datos_ph)
T_stat <- sum(log(datos_ph / x_min))

# b) Estadístico Pivotal Q (Bajo H0)
Q_cal <- 2 * beta_hipotesis * T_stat

# 3. Valores Críticos (Chi-Cuadrado)
gl <- 2 * (n - 1)
chi_inf <- qchisq(alfa / 2, df = gl)
chi_sup <- qchisq(1 - alfa / 2, df = gl)

# 4. Cálculo del P-valor (Bilateral)
# p = 2 * min(P(X < Q), P(X > Q))
p_val_lower <- pchisq(Q_cal, df = gl)
p_val_upper <- 1 - pchisq(Q_cal, df = gl)
p_valor <- 2 * min(p_val_lower, p_val_upper)

# Decisión Automática
decision <- ifelse(p_valor < alfa, "Rechazar H0", "No Rechazar H0")

# 5. Reporte de Resultados (Estilo limpio)
cat(
  "PRUEBA DE HIPÓTESIS PARA BETA (H0: B=3)\n\n",
  "Estadístico T: ", round(T_stat, 4), "\n",
  "Estadístico Q (Calc): ", round(Q_cal, 4), "\n",
  "Grados de Libertad: ", gl, "\n\n",
  "Valor Crítico Inf: ", round(chi_inf, 4), "\n",
```

```

    "Valor Crítico Sup: ", round(chi_sup, 4), "\n\n",
    "P-valor: ", format(p_valor, scientific = FALSE), "\n",
    "Decisión al 5%: ", decision, "\n"
)

```

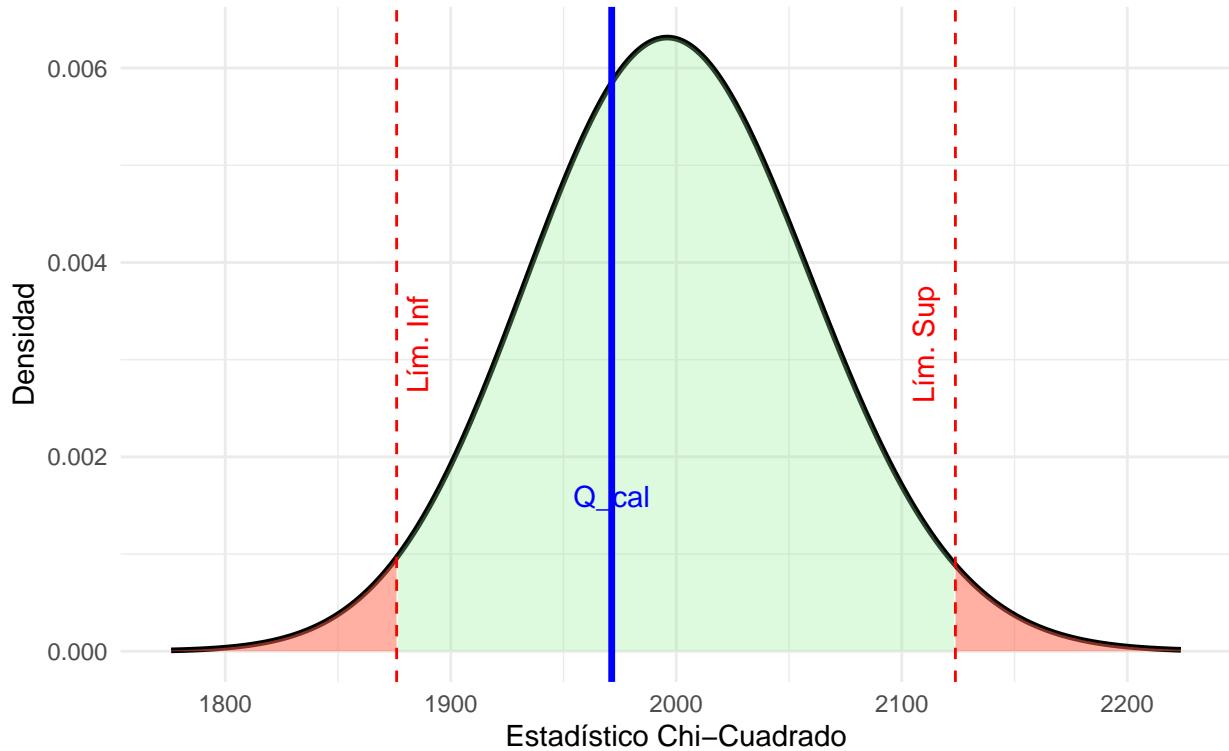
```

## PRUEBA DE HIPÓTESIS PARA BETA (H0: B=3)
##
## Estadístico T: 328.563
## Estadístico Q (Calc): 1971.378
## Grados de Libertad: 1998
##
## Valor Crítico Inf: 1876.008
## Valor Crítico Sup: 2123.78
##
## P-valor: 0.6799938
## Decisión al 5%: No Rechazar H0

```

Prueba de Hipótesis Bilateral para Beta

El estadístico calculado $Q = 1971.38$ cae en la zona de No Rechazo



4.5.5 Interpretación de la Prueba

Basándonos en los resultados obtenidos con la muestra simulada ($n = 1000$):

- Evidencia Estadística:** El estadístico de prueba calculado ($Q_{cal} \approx 1971.38$) se encuentra dentro de la región de aceptación definida por los valores críticos [1876.01, 2123.78].
- P-valor:** El P-valor obtenido (0.68) es muy superior al nivel de significancia ($\alpha = 0.05$).

Conclusión:

No existe evidencia suficiente para rechazar la hipótesis nula $H_0 : \beta = 3$. Esto confirma que el estimador EIVUM y los datos generados son consistentes con el verdadero parámetro poblacional utilizado en la simulación. Gráficamente, observamos claramente cómo la línea azul (nuestro estadístico) cae cómodamente en la zona verde (zona de no rechazo).

6. Pregunta 6 - Aplicación de una prueba de Distribución asintótica de la razón de verosimilitud.

Sea X_1, \dots, X_n una muestra aleatoria i.i.d. proveniente de una distribución Pareto(α, β), cuya función de densidad es:

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}} \mathbf{1}_{(\alpha, \infty)}(x).$$

Se desea aplicar una prueba de hipótesis basada en la distribución asintótica de la razón de verosimilitud, conforme al Teorema de Wilks.

Los momentos poblacionales están dados por:

$$E(X) = \frac{\alpha\beta}{\beta-1}, \quad \text{Var}(X) = \left(\frac{\alpha}{\beta-1}\right)^2 \frac{\beta}{\beta-2}.$$

6.1 Justificación del parámetro elegido

La prueba de razón de verosimilitud se aplica al parámetro de forma β por las siguientes razones.

El parámetro β aparece de forma **regular y diferenciable** en la función de verosimilitud, lo que permite obtener un **estimador de máxima verosimilitud interior al espacio paramétrico**, dado por:

$$\hat{\beta} = \frac{n}{\sum_{i=1}^n \ln\left(\frac{X_i}{X_{(1)}}\right)}.$$

Bajo condiciones de regularidad estándar, el **Teorema de Wilks** garantiza que el estadístico de razón de verosimilitud asociado a β converge en distribución a una chi-cuadrado con un grado de libertad:

$$-2 \ln \Lambda \xrightarrow{d} \chi_{(1)}^2, \quad n \rightarrow \infty.$$

En contraste, el estimador de máxima verosimilitud del parámetro de escala α está dado por:

$$\hat{\alpha} = X_{(1)},$$

el cual depende directamente del **extremo inferior del soporte** de la distribución. Esta característica **viola las condiciones de regularidad** requeridas por el Teorema de Wilks, por lo que una prueba de razón de verosimilitud asintótica estándar **no es apropiada para el parámetro α** .

Por estas razones, la aplicación de la prueba de razón de verosimilitud sobre el parámetro β es **teóricamente válida y estadísticamente correcta**.

6.2 Formulación de hipótesis

Se plantea la prueba bilateral:

$$H_0 : \beta = \beta_0 \quad \text{vs} \quad H_1 : \beta \neq \beta_0,$$

donde $\beta_0 = 3$.

6.3 Función de verosimilitud

La función de verosimilitud para una muestra Pareto es:

$$L(\alpha, \beta) = \prod_{i=1}^n \frac{\beta \alpha^\beta}{x_i^{\beta+1}} \mathbf{1}_{(\alpha, \infty)}(x_i).$$

6.4 Estimadores de Máxima Verosimilitud

Bajo maximización conjunta, los estimadores de máxima verosimilitud son:

$$\hat{\alpha} = X_{(1)},$$

$$\hat{\beta} = \frac{n}{\sum_{i=1}^n \ln\left(\frac{X_i}{X_{(1)}}\right)}.$$

Definimos la estadística suficiente:

$$T = \sum_{i=1}^n \ln\left(\frac{X_i}{X_{(1)}}\right).$$

```
x_min <- min(reclamos)
T_stat <- sum(log(reclamos / x_min))
T_stat
```

```
## [1] 328.563
```

6.5 Razón de Verosimilitud

La razón de verosimilitud se define como:

$$\Lambda = \frac{L(\hat{\alpha}_0, \beta_0)}{L(\hat{\alpha}, \hat{\beta})}.$$

El estadístico de prueba es:

$$-2 \ln \Lambda = 2n \left[\ln\left(\frac{\hat{\beta}}{\beta_0}\right) + \frac{\beta_0}{\hat{\beta}} - 1 \right].$$

```

beta_hat <- n / T_stat
beta_0 <- 3

LR_stat <- 2 * n * (log(beta_hat / beta_0) + beta_0 / beta_hat - 1)
LR_stat

## [1] 0.2067828

```

6.6 Distribución asintótica (Teorema de Wilks)

Bajo condiciones regulares y suponiendo verdadera la hipótesis nula, el Teorema de Wilks establece que:

$$-2 \ln \Lambda \xrightarrow{d} \chi^2_{(1)}, \quad n \rightarrow \infty.$$

6.7 P-valor y Regla de decisión

El p-valor de la prueba está dado por:

$$\text{p-valor} = P(\chi^2_{(1)} \geq -2 \ln \Lambda).$$

```

p_value <- 1 - pchisq(LR_stat, df = 1)
p_value

```

```
## [1] 0.6493008
```

Para un nivel de significancia $\alpha = 0.05$:

Rechazar H_0 si p-valor $< \alpha$.

Interpretación

Dado que:

$$\text{p-valor} = 0.6493008 > 0.05,$$

no se rechaza la hipótesis nula al nivel de significancia del 5%.

Por lo tanto, **no existe evidencia estadística suficiente** para afirmar que el parámetro de forma β sea distinto de 3.

En consecuencia, se concluye que los datos son compatibles con:

$$\beta = 3.$$

6.8 Conclusiones de la prueba

- El estadístico de razón de verosimilitud no es significativo
- La diferencia entre $\hat{\beta}$ y $\beta_0 = 3$ puede explicarse por variabilidad muestral
- El modelo bajo H_0 no es rechazado
- La prueba basada en la razón de verosimilitud confirma que el modelo bajo la hipótesis nula es consistente con los datos observados.
- Este procedimiento es válido aun cuando la población no es normal, ya que se fundamenta en la convergencia asintótica garantizada por el Teorema de Wilks.